AD_____

Award Number: **W81XWH-05-2-0027**

TITLE: **IMPACT** (Imaging and Molecular Markers for Patients with Lung Cancer: Approaches with Molecular Targets and Complementary, Innovative Treatments and Therapeutic Modalities)

PRINCIPAL INVESTIGATOR:
Waun Ki Hong, M.D.
Roy Herbst, M.D.,Ph.D.

CONTRACTING ORGANIZATION:
The University of Texas M. D. Anderson Cancer Center

Houston, TX 77030

REPORT DATE: February 2010

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
               Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

     X Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01-02-2010	Annual Report	1 FEB 2009 – 31 JAN 2010

4. TITLE AND SUBTITLE
IMPACT (Imaging and Molecular Markers for Patients with Lung Cancer: Approaches with Molecular Targets and Complementary, Innovative Treatments and Therapeutic Modalities)

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-05-2-0027

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Waun Ki Hong, M.D.

Roy Herbst, M.D., Ph.D.

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The University of Texas M.D. Anderson Cancer Center
Houston, TX 77030

Email: whong@mdanderson.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The projects in this proposal specifically target several signal transduction pathways kno wn to be critical for NSCLC pathogenesis including the EGFR pathway and the more downstream ras/raf/Mek/ERK pat hway. Th ese projects combi ne target ed approa ches usin g molecul ar an d imaging techniques to validate activity against a target and monitor response using imaging modalities specific to the receptor using either small molecules or targeted peptide approaches.

15. SUBJECT TERMS
Lung cancer, molecular markers, molecular imaging, targeted therapy

16. SECURITY CLASSIFICATION OF: Unclassified			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT U	**b. ABSTRACT** U	**c. THIS PAGE** U	UU	198	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

Appendix   (Publications)

IMPACT: Imaging and **M**olecular Markers for **P**atients with **L**ung **C**ancer: **A**pproaches with Molecular Targets, **C**omplementary, Innovative and **T**herapeutic Modalities

INTRODUCTION

Lung cancer is the most prevalent cancer worldwide and the leading cause of cancer-related mortality in both men and women in the United States. Conventional multimodality therapies (surgery, radiation and chemotherapy) have reached a therapeutic ceiling in improving the five-year overall survival rate of non-small cell lung cancer (NSCLC) patients, clinically in large part due to chemo- and radiation-resistant locoregional and metastatic spread but ultimately due to poor understanding of the disease and its resistance to the therapy.

Lung cancer is a heterogeneous disease, resulting from accumulated genetic abnormalities over years, which thus requires a coordinated attack in a truly integrated fashion on multiple altered signal pathways.   Emerging targeted therapy aims to target key molecular abnormalities in cancer and has succeeded in some tumor types such as chronic myeloid leukemia (CML) (Druker et al., 2004; Druker and Sawyers et al., 2001; Druker and Talpaz et al., 2001), gastrointestinal stromal tumor (Demetri et al., 2002), colon cancer (Hurwitz et al., 2003), and breast cancer (Howell et al., 2005). Thus, the incorporation of targeted therapy into conventional treatments appears to be a new promising approach to treatment of lung cancer.

The program project IMPACT has proposed to integrate targeted therapy in the lung cancer research program when initial clinical results showed disappointing response rates and survival benefit of epidermal growth factor receptor (EGFR) inhibitor gefitinib (Iressa™) for non-selected lung cancer patients (Herbst et al., 2002, 2003, 2004; Herbst, 2004; Kris et al., 2003; Giaccone et al., 2004). It aims to validate molecular mechanisms of targeted agents alone and in combination with chemo- and/or radiation therapies in preclinical and clinical settings. It also aims to develop effective molecular imaging and cancer cell-targeted peptide-based delivery tools to help improve efficacy of the targeted agents. Specifically, our objectives are:

- To validate preclinically and clinically several key signaling pathways and their agents for therapeutic potentials alone or in combination with each other or with chemo and/or radiotherapy
- To explore applications of molecular imaging for targeted therapy and identify cancer cell-targeted peptides for systemic delivery of therapeutic and imaging agents
- To discover and evaluate new molecular abnormalities and therapeutic predictors in lung cancer
- To develop an educational program for teens and young adults for smoking risk and resultant lung cancer occurrence.

IMPACT is composed of 6 research projects, 1 Biostatistics Core, 1 Molecular Pathology Core, 1 Molecular Imaging Core, 2 career development projects, and 2 developmental research projects. Here we present their scientific progresses in the fourth grant year as follows.

Project 1: Targeting epidermal growth factor receptor signaling to enhance response of lung cancer to therapeutic radiation.

(PI and co-PI: Raymond E. Meyn, Ph.D., Ritsuko Komaki, M.D.)

In spite of significant technical advances including intensity-modulated radiation therapy (IMRT) and chemoradiation, locally advanced lung cancer continues to have a dismal prognosis as many patients' tumors appear to be resistant to radiation therapy. The molecular basis for radiation resistance is not fully understood, but tumor cells have an enhanced survival response that involves increased capacity for DNA repair and suppressed apoptosis. Both apoptosis propensity and DNA repair capacity are thought to be partly controlled by the upstream signal transduction pathways triggered by EGFR activation, which is constitutively activated in many NSCLCs, and its activation leads to a radiation-resistant phenotype. We hypothesize that the response of NSCLC to radiation can be improved through the use of inhibitors of EGFR signaling.

Aim 1   To test the combination of external beam radiation and the selective EGFR-tyrosine kinase inhibitor erlotinib (Tarceva) in locally advanced NSCLC.

<u>Summary of Research Findings</u>
This trial uses chemoradiotherapy followed by a molecularly targeted treatment (erlotinib, targeting the EGFR tyrosine kinase) with/without radiotherapy for stage III NSCLC to improve the therapeutic ratio (i.e., increase malignant cell cytotoxicity without increasing normal cell cytotoxicity). The primary objective is to determine the efficacy of concurrent chemoradiotherapy followed by erlotinib/radiation and erlotinib alone, which will be repeated for 7 weeks, then two cycles of consolidation chemotherapy as measured by progression-free survival. Secondary objectives include determining: 1) safety, as measured by the rate of grade 3 or worse non-hematological toxicity (dose-limiting toxicity, DLT) occurring prior to the beginning of consolidation therapy (including all toxicities attributed to chemoradiation occurring within 90 days of the start of radiation therapy); 2) compliance, which is defined to be completion of concurrent chemoradiation and erlotinib/radiotherapy with no more than minor variations; 3) response rate (complete and partial response rates); 4) overall survival rates (one- and two-year rates, median survival); 5) disease (local) control rate; 6) association between EFGR expression and toxicity, response, overall survival, and progression (exploratory analysis); and 7) association between EGFR expression and response correlated with imaging study. A total of 48 patients will be treated with concurrent chemoradiation [RT: 63 Gy/35 fractions/7weeks (+/- 5 days), 1.8 Gy/fraction, a total dose of 63.0 Gy in 35 fractions over 7 weeks plus paclitaxel, 45mg/m², and carboplatin, AUC=2, weekly on day 1 for 7 weeks] and erlotinib (erlotinib, 150 mg p.o. daily for 7 weeks, starting with radiotherapy on day 2-5 followed by erlotinib 150 mg p.o. alone on day 6-7 for 7 weeks). Patients will get one month off treatment, followed by consolidation therapy of chemotherapy alone (weeks 11-17: paclitaxel, 200 mg/m², and carboplatin, AUC=6, every 3 weeks for two cycles).

Currently, a total of 38 patients were enrolled on this protocol between November 2007 and February, 2010, with an average accrual rate of 1.5 patients/month. The accrual rate for the first several months after activation was lower than predicted, but has increased to 2 patients/month since August 2008. We anticipate completion of the study within the next 8 months during the requested no-cost extension of our grant.

Patient Status
As of February 2010, two patients were discontinued from the protocol treatment: one patient due to disease progression during the treatment, the other patient discontinuation was due to acute MI.

Patients screened: 55
Screening failures: 17
Patients enrolled to the protocol: 38
Patients completed treatment according to the protocol: 30
Patients still on concurrent chemoradiation: 2
Patients still on consolidation chemo: 4
Patients who could not complete Tarceva regimen according to the protocol: 2
Patients who missed one or two consolidation chemo: 7

Local recurrence (within the Radiation field): 3
Regional recurrence (Outside Radiation field): 1

Distant metastasis: 7 (3 liver mets, 1 brain mets, 1 right lower lobe mets, 1 T2 bone mets, 1
   opposite side of the neck mets)
Possible L2 metastasis (Not biopsied): 1

Five patients died during follow-up in the clinical trial: #1 one had heart attack; #2 had major organ failure; #3 had neutropenic fever and pneumonia; #4 had abdominal carcinimatosis from non small cell lung cancer; #5 had pulmonary embolism, radiation pneumonitis, fungal infection, and respiratory failure. Three of these patients also had liver metastasis.

Response
Thirty patients have completed the treatment regimen per protocol (Figure 1). Response by RECIST criteria for the patients accrued to date is under analysis, and will be reported upon completion of the trial. It should be noted that per RECIST, tumor response is measured by CT; however, chemoradiotherapy causes fibrotic changes in the lung, which makes accurate measurement of the residual tumor impossible. Therefore, metabolic response by PET may be more accurate after chemoradiotherapy.

As an exploratory analysis, twenty-eight of 30 patients were evaluated for a PET SUV measurable response to the treatment; fourteen patients (47%) had a complete response (CR); thirteen patients (43%) achieved partial response (PR) and one patient achieved stable disease (SD). The total tumor response rate was 90% (27/30) by PET. Two patients did not have post-chemoradiotherapy PET report due to their death before the scheduled follow-up PET exam.
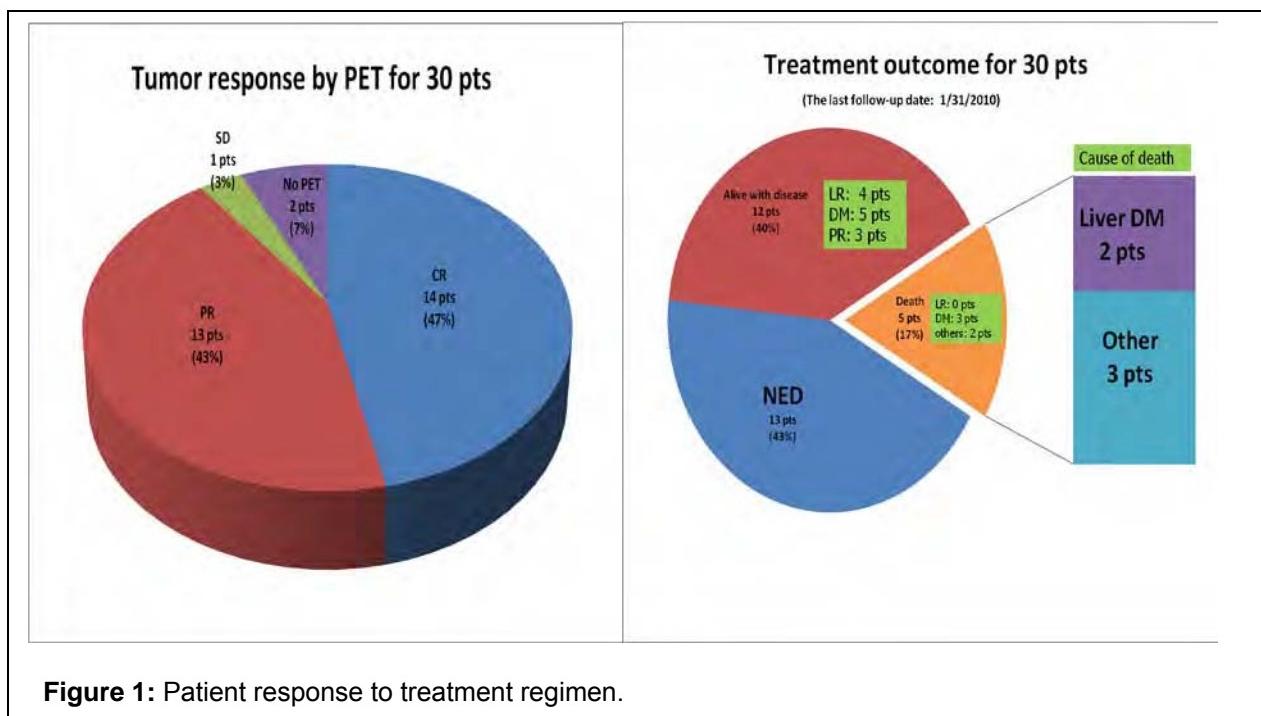
Figure 1: Patient response to treatment regimen.

Toxicity

Toxicity dat a i s av ailable for 38 pat ients who hav e either c ompleted ther apy or ar e pr esently receiving treatment (Figure 2). Severe acute toxicities (Grade 3 or higher according to C TC.3) related to treatments were recorded as the following events:

- Rash, Grade 3 in 4 patients
- Acne, Grade 2 in 12 patients
- Acne, Grade 3 in 2 patients
- Diarrhea, Grade 2 in 4 patients
- Diarrhea, Grade 3 in 2 patient
- Pneumonitis, Grade 3 in 2 patients; pneumonitis Grade 5 in 1 patient
- Leukopenia, Grade 3 in 12 patients, Grade 4 in 1 patient
- Neutropenia Grade 3 in 5 patient, Grade 4 in 2 patients
- Thrombocytopenia, Grade 3 in 1 patient
- Hypomagnesemia, Grade 3 in 1 patient
- Hypokalemia, Grade 3 in 2 patient
- Pneumonia, Grade 3 in 7 patients
- Dehydration, Grade 3 in 3 patients.

Protocol 2005 -1023 appear s to hav e a toxicity pr ofile that i s c omparable to other chemoradiotherapy r egimens r eported i n thi s patient pop ulation. P reliminary r esponse data seems fav orable, al though i t i s to o ear ly i n the s tudy to be c ertain of w hether thi s r esult i s significant. Continuation of the study is warranted.

Correlation of bi omarkers fr om the pati ent bi opsy and bl ood s pecimens w ith r esponse i s underway. We will determine if there are any correlations between biomarkers and response as well as with toxicity w hen 48 pati ents c omplete tr eatment. The r esults w ill be s ubmitted to a

major international scientific meeting such as AACR or others. If the rate of accrual of 2 patients/month continues, we will complete this trial by the end of this year.

Summary:
1. ID 2005-1023 has shown 90% response (CR 47% and PR43%).
2. The rate of 7.9% (3/38) Grade 3-5 pneumonitis is acceptable.
3. One treatment-related grade 5 toxicity (pneumonitis) has occurred to date.

Analysis of IHC and plasma will be done after completion of this study late this year.

Aim 2 **To test the hypothesis that activation of the EGFR pathway leads to radiation resistance in NSCLC cells due to an enhanced capacity for repairing DNA lesions.**

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 3 **To test the hypothesis that clinically useful inhibitors of EGFR signaling abrogate DNA repair capacity, restore apoptotic response and radiosensitize NSCLC cells.**

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 4 **To test the hypothesis that targeting both EGFR and its downstream signaling pathways will have at least an additive radiosensitizing effect on NSCLC.**

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 5 **To test whether the strategies developed in Specific Aims 2-4 have efficacy in a xenograft tumor model.**

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Key Research Accomplishments

- Enrolled 38 patients onto the erlotinib (Tarceva) plus radiotherapy for locally advanced NSCLC trial, and completed evaluation of 30 of these patients with excellent response and minimal toxicity.
- Discovered the r elationship bet ween the epithelial-to-mesenchymal tr ansition and radiosensitivity of NSCLC cells (preclinical data).
- Demonstrated that pr etreatment w ith gefi tinib e xerts a r adioprotection of H 1299-CDH1 cells.
- Demonstrated that s mall m olecule i nhibitors of both c -Met and IG F-1R pr oduce a significant radiosensitizing effect on NSCLC cells.
- Completed an as sessment of the c ombination of er lotinib (Tarceva) and r adiation i n a NSCLC xenograft tumor model.

Conclusions

We c onclude that the epithelial-to-mesenchymal tr ansition (EMT) pl ays a s ignificant r ole i n governing not j ust the intrinsic r adiosensitity of N SCLC c ells, but also thei r s ensitivity to inhibitors of the epi dermal gr owth factor r eceptor (EGFR) and the abi lity of s uch i nhibitors to radiosensitize these cells. It w ould be us eful to assess the EMT status of pati ents treated with these c ombinations. In spite of thi s fi nding, r esults suggest that such combinations m ight be useful in the clinic. In addition, we conclude that targeting other growth factor receptors such as the c-Met and IG-F1R receptors may be an alternative strategy to using EGFR inhibitors.

Project 2: Molecular Imaging of EGFR Expression and Activity in Targeting Therapy of Lung Cancer

(PI and co-PI: Juri Gelovani, M.D., Ph.D.; Roy Herbst, M.D., Ph.D.)

Aim 1 **To synthesize novel pharmacokinetically optimized ^{124}I and ^{18}F-labeled IPQA derivatives for PET imaging of EGFR kinase activity and conduct *in vitro* radiotracer accumulation studies in tumor cells expressing different levels of EGFR activity.**

This aim was completed and summarized in the previous reports.

Aim 2 **To assess the biodistribution (PK/PD) and tumor targeting by novel ^{124}I and ^{18}F-labeled EGFR kinase-specific IPQA derivatives using PET imaging in orthotopic mouse models of lung cancer and compare *in vivo* radiotracer uptake/retention with phospho-EGFR levels *in situ*.**

Summary of Research Findings

Differential sensitivity of NSCLC cell lines to EGFR kinase inhibition.
In vitro cell growth of different NSCLC cell lines was differentially inhibited by gefitinib (Iressa™) in a dose-dependent manner (Figure 2A). The most gefitinib-sensitive cell lines were the H3255 cells expressing L858R mutant EGFR, whereas the H1975 cells expressing both the L858R and T790M EGFR mutations were significantly more resistant. The H441 and PC14 cells expressing wild-type EGFR exhibited significant resistance to gefitinib.

Differential expression of EGFR in NSCLC cell lines.
The level of total EGFR expression under serum-starved conditions as measured by ELISA was similar in H3255 and H441 cells and significantly higher than levels in PC14 and H1975 cells (Figure 2B).

Preferential accumulation of ^{18}F PEG6-IPQA in NSCLC cells with L858R EGFR mutation.
All four cell l ines demonstrated a r apid uptake of ^{18}F-PEG6-IPQA (Figure 2C) during the i nitial phase (first 20 m in). T hereafter, the accumulation of ^{18}F-PEG6-IPQA r eached a pl ateau i n H441, H1975, and PC14 cells at about 30-40 cells/medium concentration ratio. In contrast, in H3255 cells the accumulation of ^{18}F-PEG6-IPQA continued to increase up to one hour and thereafter r eached a pl ateau at c ells/medium c oncentration r atio of 400 -600 (Figure 2D). The magnitude of ^{18}F-PEG6-IPQA accumulation in H3255 cells was more than 10-fold higher than in H441 cells, despite the similarity in levels of EGFR expression in the two cell lines, as measured by E LISA (Figure 2B). The "washout" s tudy dem onstrated a s ignificant r etention of ^{18}F-PEG6-IPQA in H3255 cells (at about 400 cell/medium concentration ratio), which accounted for about 65% of total accumulated radioactivity between 60 and 120 min of incubation.

In contrast, almost 50% of ^{18}F-PEG6-IPQA radioactivity could be washed out fr om H 441 cells, more than 60% from PC14 c ells, and m ore than 70% from H1975 cells. After washout in PC14 and H1975 cells, the passive volume of ^{18}F-PEG6-IPQA di stribution was less than 10. The magnitude of ^{18}F-PEG6-IPQA accumulation in all tested c ell l ines w as significantly dec reased i n the presence of gefitinib at 100 µM/L in the c ulture medium.
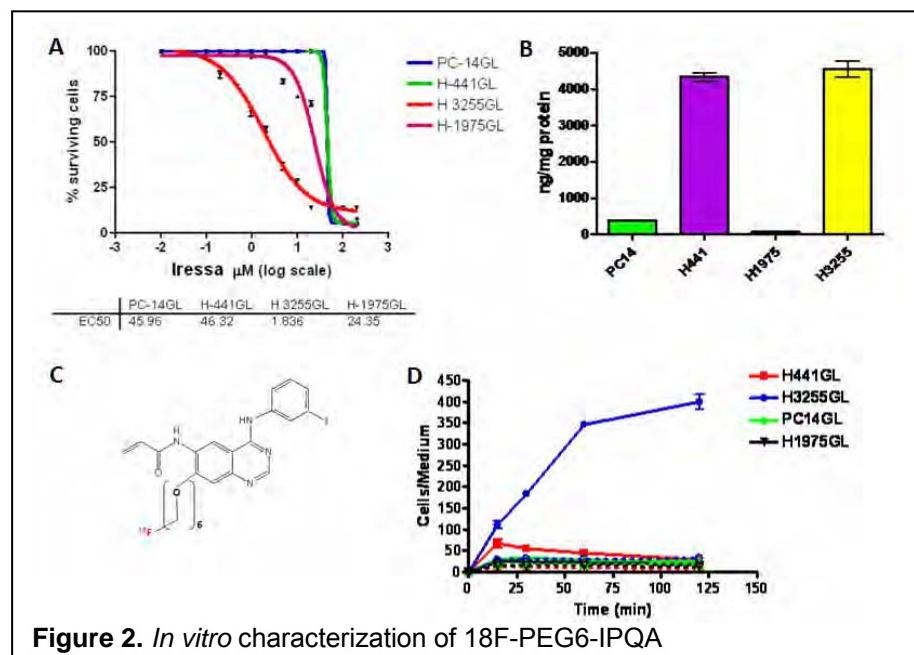
Figure 2. *In vitro* characterization of 18F-PEG6-IPQA

Preferential and irreversible binding of ^{18}F-PEG6-IPQA to the L858R mutant EGFR kinase. The autoradiograhic and Western bl ot anal ysis of el ectrophorograms of pr otein ex tracts fr om different NSCLC cells incubated with ^{18}F-PEG6-IPQA demonstrated preferential and irreversible covalent bi nding of ^{18}F-PEG6-IPQA to L858R mutant E GFR k inase do main, c orresponding to 172 kD protein band (Figure 3). In H441 cells, the irreversible binding to wild-type EGFR kinase was s ignificantly l ower than in H 3255 c ells.

Figure 3. *In vitro* characterization of 18F -PEG6-IPQA

This obs ervation bec ame m ore evident when comparing the i ntensity of r adioactive bands (Figure 3 A) and c orresponding Western B lot bands (Figure 3B) of d ifferent dilutions of c ellular p rotein ex tracts. In particular, t he 1:10 di lution of H 441 ex tract contained significantly m ore EGFR protein than the 1:10 di luted ex tract fr om H 3255 cells; how ever, the i ntensity of the radioactive band c orresponding to 1:10 dilution of H3255 c ells w as s ignificantly higher (at least a l og order) than th at in the 1:10 di luted ex tract of H 441 c ells. Autoradiographic det ection w as m ore sensitive than i mmunoblotting and chemiluminescent dete ction, as e videnced by the pr esence of a f aint r adioactive band in 1:1000 di luted ex tracts of H 3255 c ells, while the c orresponding band on Western blot i s not detec table. The i rreversible covalent binding of ^{18}F-PEG6-IPQA to EGFR kinase domain was barely detectable in undi luted ex tracts of P C14 c ells expressing l ow l evels of w ild-type E GFR and i n undi luted ex tracts of H 1975 c ells

expressing L858R/T790M dual mutant EGFR. More importantly, the enzyme inhibition studies demonstrated that cold (non-radiolabeled) F-PEG6-IPQA inhibits the recombinant L858R EGFR kinase a log order better than the recombinant wild-type EGFR kinase, and almost 50-fold better than the L858R/T790M dual mutant EGFR kinase (Figure 3C). Together with the results of *in vitro* radiotracer accumulation studies, these data confirm the selectivity of ^{18}F-PEG6-IPQA to active mutant L858R EGFR kinase.

Aim 3 **Using selected ^{124}I or ^{18}F-labeled IPQA derivative, to conduct pre-clinical studies in animals with orthotopic models of lung cancer xenografts with different levels of EGFR expression/activity, and to assess the value of PET imaging as the inclusion criterion for therapy by EGFR inhibitors, as well as for monitoring the efficacy of treatment with EGFR-targeted drugs.**

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 4 **Perform pilot clinical PET imaging studies with the optimized ^{124}I or ^{18}F-labeled IPQA derivative under the RDRC guidelines in patients with NSCLC undergoing adjuvant therapy before tumor resection or biopsy. Compare PET image-based measures of EGFR activity with immunohistochemical measures of phospho-EGFR *in situ*.**

Summary of Research Findings
In vivo PET/CT Imaging with ^{18}F-PEG6-IPQA enables detection of NSCLC expressing L858R mutant EGFR.

In vivo dynamic PET/CT imaging demonstrated a rapid accumulation of ^{18}F-PEG6-IPQA in H3255 and, to a lower degree, in H441 subcutaneous (s.c.) tumor xenografts (Figure 4A). The ^{18}F-PEG6-IPQA accumulation at 120 post i.v. injection reached 2.34±0.13 %ID/g in H3255 tumors and 1.59±0.44 %ID/g in H441 tumors. Tumor-to-muscle ratios of ^{18}F-PEG6-IPQA accumulation for H3255 xenografts had reached the level of 2.08±0.19 at 120 min post i.v. injection, and at the level of 1.47±0.08 for H441 tumor xenografts (Figure 4C). In contrast, the level of ^{18}F-PEG6-IPQA accumulation was insignificant in PC14 and H1975 s.c. tumor xenografts reaching 0.90±0.11 and 1.05±0.09 %ID/g in PC14 and H1975 tumors, respectively (Figure 4A), which was similar to the 0.94±0.22 %ID/g level observed in the muscle (reference tissue). Thus, the magnitude of ^{18}F-PEG6-IPQA accumulation in H3255 was 1.4, 6.3, and 7.6 fold more ($p < 0.001$) than H441, PC14, and H1975, respectively. Tumor-to-muscle ratios of ^{18}F-PEG6-IPQA accumulation in PC14 and H1975 tumor xenografts reached equilibrium of about 1 starting at 20 minutes post i.v. injection (Figure 4C). The unidirectional rate of ^{18}F-PEG6-IPQA accumulation (*Ki*), calculated using Patlak graphical analysis approach, was significantly higher in H3255 tumor xenografts than in H441, PC14, and H1975 xenografts (p<0.05) (Figure 5A). Logan graphical analysis approach demonstrated increased binding potential (BP) of ^{18}F-PEG6-IPQA in H3255 tumors, as compared to other tumors (p<0.05) (Figure 5B).

Pharmacokinetics of ^{18}F-PEG6-IPQA in tumor-bearing mice.
After intravenous injection, the ^{18}F-PEG6-IPQA exhibited a bi-exponential kinetics of clearance, with half-lives for the rapid and slow phases of 1.15 and 21.32 min, respectively. The initial clearance of ^{18}F-PEG6-IPQA–derived radioactivity from circulation was via the hepato-billiary route; however, starting at 40 min post i.v. injection, the renal clearance of ^{18}F-PEG6-IPQA –derived radioactivity became predominant at a rate of 27.98 %ID/ml/min (Figures 4E,F).

Treatment with gefitinib significantly inhibits ^{18}F-PEG6-IPQA accumulation in NSCLC expressing L858R mutant EGFR.

In vivo dynamic P ET/CT i maging dem onstrated a significant reduction of ^{18}F-PEG6-IPQA accumulation in H3255 tumor xenografts after treatment with gefitinib (100 mg/kg 1 hour before administration of ^{18}F-PEG6-IPQA) to the level of 1.38±0.43 %ID/g and a tumor-to-muscle ratio of 1.56±0.36 at 120 min post i.v. injection (Figures 4B,D). In contrast, the level of ^{18}F-PEG6-IPQA accumulation in H441 tumor xenografts was 1.58±0.01 %ID/g with a tumor-to-muscle ratio of 1.77±0.53 at 120 min post i.v. injection (Figures 4B,D), which was similar to pre-treatment values (Figures 4A,C). In PC14 and H1975 tumor xenografts, the magnitude and tumor-to-muscle r atios of ^{18}F-PEG6-IPQA ac cumulation fur ther dec reased (Figures 4B,D). Both Patlak and Logan graphical a nalyses dem onstrated a statistically significant r eduction of ^{18}F-PEG6-IPQA accumulation rate (Ki) and BP in H3255 tumor xenografts (p<0.05), whereas the trends in reduction of ^{18}F-PEG6-IPQA ac cumulation i n other tum or x enografts w ere not s tatistically significant (Figure 5).
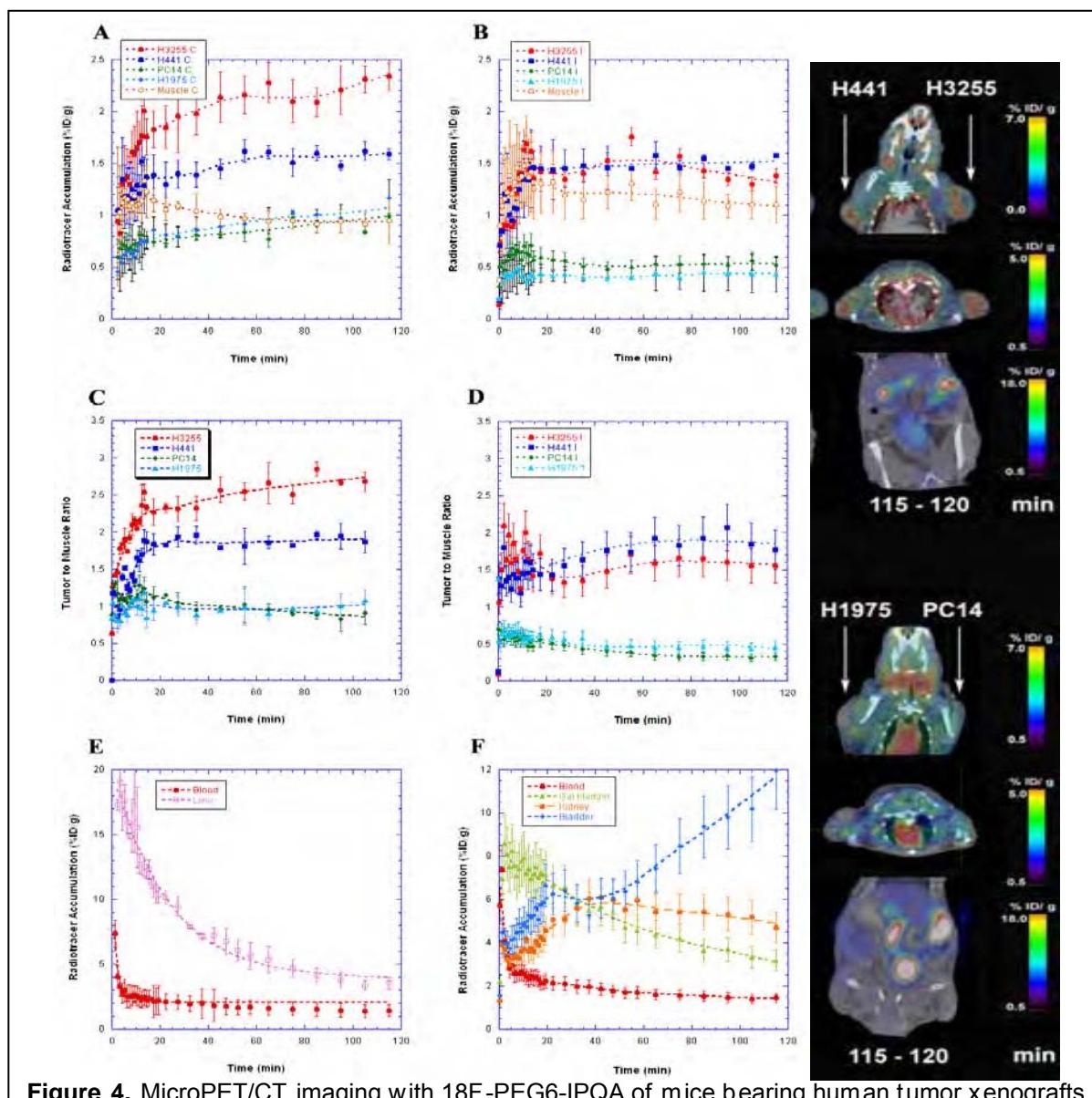
Figure 4. MicroPET/CT imaging with 18F-PEG6-IPQA of mice bearing human tumor xenografts grown from various human NSCLC cell lines expressing different EGFR mutants.
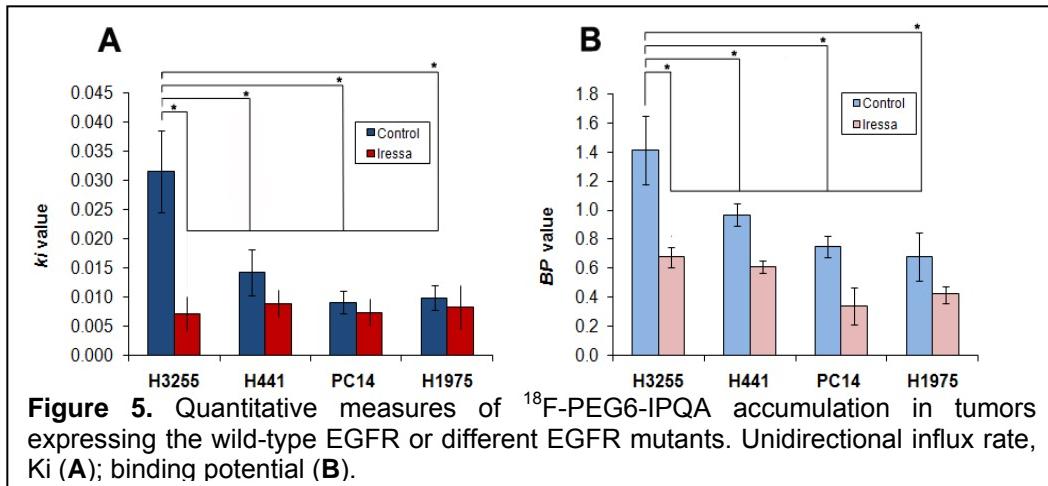
Figure 5. Quantitative measures of ^{18}F-PEG6-IPQA accumulation in tumors expressing the wild-type EGFR or different EGFR mutants. Unidirectional influx rate, Ki (**A**); binding potential (**B**).

Quantitative autoradiography confirms the results of *in vivo* PET/CT imaging with ^{18}F-PEG6-IPQA.

Quantitative autoradiography (QAR) demonstrated preferential accumulation of ^{18}F-PEG6-IPQA in viable portions of H3255 tumor xenografts (Figure 6). Also, QAR revealed intratumoral heterogeneity of ^{18}F-PEG6-IPQA. High levels of ^{18}F-PEG6-IPQA-derived radioactivity accumulated in the viable portions of H3255 tumor xenografts (3.01± 0.84 %ID/g). In contrast, no specific ^{18}F-PEG6-IPQA accumulation was observed inside the necrotic tumor areas. The accumulation of ^{18}F-PEG6-IPQA in H441 tumors (2.53±0.15 %ID/g) was lower than that in H3255 xenografts, while in H1975 (1.27±0.12 %ID/g) and PC14 (1.54±0.14 %ID/g) tumors, accumulation was significantly lower. Pretreatment with gefitinib resulted in about 4-fold decrease in ^{18}F-PEG6-IPQA accumulation in H3255 tumors (to 0.92±0.13 %ID/g) and ~3 times in H441 (0.78±0.06 %ID/g). There were almost no differences in ^{18}F-PEG6-IPQA accumulation in PC14 and H1975 tumor xenografts before and after treatment with gefitinib (1.17±0.10 and 1.05±0.01 %ID/g, respectively).
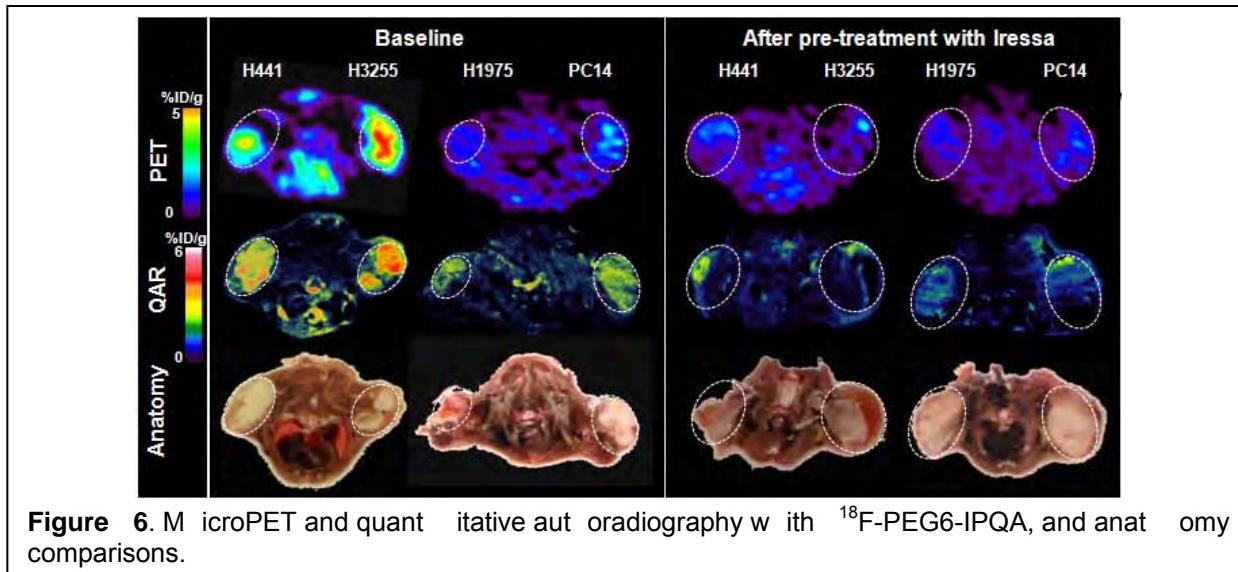
Figure 6. MicroPET and quantitative autoradiography with ^{18}F-PEG6-IPQA, and anatomy comparisons.

In summary, our xenograft studies in mice demonstrated a significantly increased accumulation of ^{18}F-PEG6-IPQA in H3255, as compared to H441, PC14, and H1975 tumor cells. Due to the similarities in the levels of L858R EGFR expression in H3255 and H441 cells, preferential

accumulation of ^{18}F-PEG6-IPQA in H3255 tumors can be explained at least in part by the presence of the L585R activating mutation in the EGFR kinase domain [1, 2, 3]. The L858R mutation dramatically enhances the magnitude of EGFR signaling activity in H3255 cells, as compared to H441 cells that over-express wild-type EGFR [1, 3] and PC14 cells that express wild-type EGFR at a low level. As demonstrated by our SAR modeling studies, the activation status of EGFR kinase appears to be very important for the binding mode of ^{18}F-PEG6-IPQA and formation of covalent bond with Cys773. A similar conclusion was drawn from the results of previous studies with structurally similar 4-anilinoquinazoline-based reversible inhibitors (gefitinib and erlotinib) that demonstrated these compounds had more than 20-fold increase in the affinity and inhibitory activity for the L858R-mutant, as compared to wild-type, EGFR kinase [4], presumably due to a shift in equilibrium toward the active state in L858R-mutant EGFR kinase.

While the H1975 cells do carry the L858R EGFR mutation that boosts EGFR signaling, the second T790M mutation causes resistance to reversible EGFR inhibitors, such as gefitinib and erlotinib, by interfering with binding of these inhibitors to the ATP-binding site of EGFR kinase [5, 6]. This explanation is supported by the results of an autoradiographic study with Western blot analyses of electrophorograms of proteins extracted from different NSCLC cells incubated *in vitro* with ^{18}F-PEG6-IPQA, which demonstrated increased irreversible binding of ^{18}F-PEG6-IPQA to the L858R-mutant EGFR kinase domain, as compared to the wild-type EGFR kinase, and no irreversible binding to L858R/T790M dual-mutant EGFR kinase.

Study of pharmacokinetics, biodistribution, radiation dosimetry, and metabolites of ^{18}F-PEG6-IPQA.

One of the most important components of an Investigational New Drug (IND) protocol for novel radiolabeled imaging agents is the assessment of Pharmacokinetics (PK), Metabolism, and Radiation Dosimetry. For this application, we have performed dynamic PET/CT imaging studies, funded from non-DoD sources, with ^{18}F-PEG6-IPQA in non-human primates (rhesus macaques; 3 males and 3 females). Preliminary data from 6 primates imaged with ^{18}F-PEG6-IPQA were collected for the purpose of obtaining initial estimates of average organ doses per MBq, as well
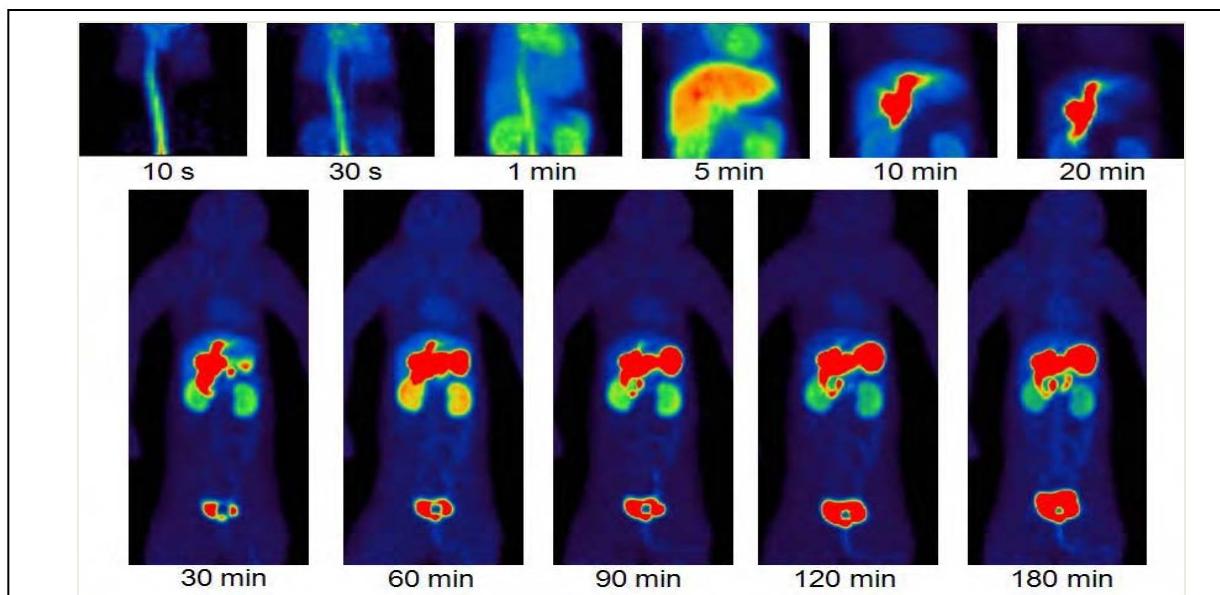
Figure 7. Dynamic PET images of ^{18}F-PEG6-IPQA –derived radioactivity distribution acquired in a rhesus monkey.

as between subject variations. Radiation-absorbed doses in 25 different organ systems were estimated for 3 female and 3 male primates, from which humanized organ doses were constructed. These data are used to estimate the two quantities relevant for the estimation of radiation dose to be experienced by patients included in this study: the population mean and the standard deviation of radiation exposure between patients for each organ. Based on these data, our clinical protocol will require that each subject in the first cohort (3 patients) receive the ^{18}F-PEG6-IPQA injection with a maximum total activity of 70 Mbq. An example of spatial and temporal dynamics of ^{18}F-PEG6-IPQA in a rhesus macaque is shown in Figures 7 and 8.
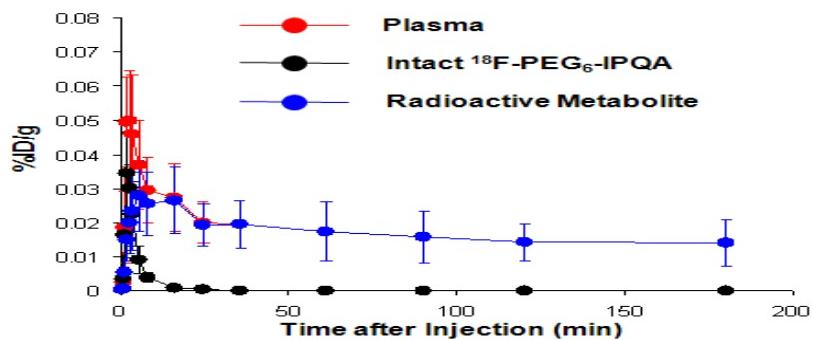
Figure 8. Time-activity curves of ^{18}F-PEG6-IPQA and major radioactive metabolite (^{18}F-PEG6) in blood plasma.

GLP extended single-dose toxicology studies in rats.

Per FDA requirements, acute toxicology studies in rats have been conducted by the Pre-Clinical Services group of the Charles River Laboratories (Spencerville, OH) in compliance with the Good Laboratory Practice (GLP) regulations as described by the FDA (21 CFR Part 58). The purpose of thi s study was to evaluate the potential toxicity of Fluoro-PEG6-IPQA when administered to rats as a single intravenous (bolus) injection, followed by a 14-day recovery period. The study design was as follows:

Group No.	No of Toxicity (Recovery) Animals		Dose Material	Target Dose Level (µg/kg)	Dose Volume (mL/kg)	Target Dose Concentration (µg/mL)	Actual Dose Level (µg/kg)
	Male	Female					
1	10 (5)	10 (5)	6% Ethanol in Saline	0	7.75	0	0
2	10 (5)	10 (5)	Fluoro-PEG6-IPQA	155.0	7.75	20	136.4[a]
[a]Based on dose formulation analysis report (AEZ00020AO-1-001-1), the concentration of formulation was approximately 2% below the acceptable specification (±10%) of nominal concentration. The Group 2 actual dose concentration is referred as 136.4 µg/kg throughout this report.							

The following parameters and end points were evaluated in this study: clinical signs, body weights, body weight changes, food consumption, clinical pathology parameters (hematology, coagulation, clinical chemistry, and urinalysis), gross necropsy findings, organ weights, and histopathologic examinations.

There was no test article-related mortality observed in this study. All toxicity and recovery phase animals survived until scheduled euthanasia. There were no test article-related clinical

observations during the dosing and recovery phases in males and females. There were no toxicologically meaningful differences in mean body weights, body weight changes, food consumption, hematology parameters, coagulation, clinical chemistry parameters, urinalysis parameters (macroscopic and microscopic), and absolute or relative organ weights in test article-treated males or females during the toxicity and recovery phases. There were no toxicologically meaningful test article-related gross necropsy findings observed at termination of the toxicity and recovery phases. There were no test article-related microscopic changes noted during the main or recovery sacrifices in males or females. In both study groups, a procedure-related hemorrhage and inflammation noted at the injection sites at Study Day 2 had essentially resolved by Study Day 15. In addition, varying numbers of abnormal spermatozoa and varying degrees of aspermia in the epididymides were noted and attributed to immaturity. These findings also resolved during the recovery phase.

We conclude that administration of Fluoro-PEG6-IPQA to rats as a single intravenous (bolus) injection was well tolerated at a dose level of 136.4 µg/kg. No systemic toxicity was observed and no target organs were identified. Based on these results, the no-observed-adverse-effect level (NOAEL) was determined to be 136.4 µg/kg in both sexes.

Clinical Protocol

The clinical protocol for a phase I study of ^{18}F-PEG6-IPQA in NSCLC patients has been fully developed by the Co-Principal Investigators, Drs. David Stewart (Department of Thoracic/Head and Neck Medical Oncology) and Don Podoloff (Department of Nuclear Medicine), and approved by the Nuclear Medicine Department at M.D. Anderson Cancer Center. The objectives of the phase I study are summarized as follows:

Primary Objectives:

- To determine the optimum dosimetry of ^{18}F-PEG6-IPQA sodium injection based on critical organ safety and detection sensitivity.
- To obtain data on ^{18}F-PEG6-IPQA distribution, pharmacokinetics and metabolites.
- To assess the safety of a single intravenous administration of ^{18}F-PEG6-IPQA in subjects with solid tumors.

Secondary Objectives:

- To obtain preliminary data on the feasibility of detection of both primary and metastatic tumor lesions using ^{18}F-PEG6-IPQA PET as compared to standard of care modalities [e.g., ^{18}F-fluorodeoxyglucose PET, contrast enhanced static computed tomography (CECT), magnetic resonance imaging (MRI), bone scintigraphy and/or US].
- To correlate the magnitude of tumor uptake and retention of ^{18}F-PEG6-IPQA with tumor EGFR expression and/or drug response.

The rationale of the current study will be to determine the optimum dosimetry of this novel PET imaging agent in patients with NSCLC; and to demonstrate that it is feasible and safe to be employed in routine PET imaging use based on characteristics of its biodistribution, pharmacokinetics, and metabolites. It is also aimed to assess the detectability of this agent for NSCLC and to evaluate the correlation of magnitude of ^{18}F-PEG6-IPQA retention in tumor with EGFR expression and/or drug response.

Eligibility criteria for this study have been carefully considered to ensure the safety of the study subjects and to ensure that the results of the study can be used. It is imperative that subjects fully meet all eligibility criteria. For entry into the study, the following criteria must be met.

Inclusion Criteria
1. All patients must give written informed consent.
2. Patients should have pathologically or cytologically confirmed non-small cell lung cancer with clinical or radiological evidence that it is not amenable to therapy with curative intent.
3. Patients should be potential candidates for therapy with an EGFR tyrosine kinase inhibitor or with an anti-EGFR monoclonal antibody by clinical criteria.
4. Patients should have clinical characteristics that would suggest an increased probability of benefit from an EGFR inhibitor. Specifically, they should have either:
    - Less than a 10 pack-year smoking history AND a latency period from last tobacco use to diagnosis of longer than 10 years
    - AND either lung adenocarcinoma or NSCLC not otherwise specified, OR
    - Known EGFR mutations OR high EGFR gene copy number
5. Patients should have at least one tumor deposit that is > 1.0 cm in diameter, and that is amenable to imaging.
6. Patients should be ECOG performance status 0-2.
7. Patients with brain metastases are eligible provided they meet all other eligibility criteria and do not require corticosteroids or enzyme-inducing anticonvulsants and provided it is felt clinically that they will not require radiotherapy in the three (3) weeks subsequent to their participation in the study.
8. Women of childbearing potential must agree to use adequate contraception (hormonal or barrier method of birth control; abstinence) prior to study entry and for the duration of study participation. Childbearing potential will be defined as women who have had menses within the past 12 months, who have not had tubal ligation or bilateral oophorectomy. Should a woman become pregnant or suspect that she is pregnant while participating in this study, she should inform her treating physician immediately. The patient, if a man, agrees to use effective contraception or abstinence.

Exclusion Criteria

1. Prior therapy with an EGFR inhibitor or an anti-EGFR monoclonal antibody.
2. Radiotherapy, chemotherapy or any investigational agent within the previous 4 weeks.
3. A non-investigational targeted agent within the previous 2 weeks.
4. Thoracic or abdominal surgery within the previous 2 weeks.
5. A tumor that is known to have a KRAS mutation.
6. Squamous cell, large cell undifferentiated, neuroendocrine or small cell undifferentiated carcinoma of the lung.
7. A known other currently active malignancy (benign tumors and benign polyps, basal cell carcinomas of skin, superficial papillary bladder tumors, and pre-invasive carcinoma of the cervix are permitted).
8. Physical inability to undergo a scanning procedure (e.g., inability to lie flat for the required period of time – three sessions of roughly an hour each with ten minutes' rest in between).
9. Serum creatinine >1.5 x ULN, bilirubin >1.5 x ULN, AST > 3 x ULN.
10. Hemoglobin < 8 g/dL, absolute neutrophil count < 1,500/mm^3, platelet count <100,000/mm^3.
11. Potentially life-threatening arrhythmia; myocardial infarct within the previous 3 months; unstable angina, or angina at rest; congestive heart failure (New York Heart Association Functional Classification class II or worse), uncontrolled hypertension (systolic BP > 160 or diastolic BP >100).

12. Active acute infection (i.e. currently treated with antibiotics). Patients with chronic infections such as hepatitis B or C, mycobacterium avium or similar infections will be eligible provided they meet all other eligibility criteria.
13. Oxygen saturation <90% on room air.
14. Clinical requirement for systemic corticosteroids for control of cerebral edema or for enzyme-inducing anticonvulsants. (Inhaled steroids and systemic steroids for COPD are permitted).
15. Pregnant or nursing.
16. Any condition that is unstable or could jeopardize the safety of the patient and his or her compliance in the study, in the investigator's judgment.

Up to 15 evaluable subjects are planned to be included at the single study center M.D. Anderson Cancer Center. Subjects are considered evaluable if they undergo administration of ^{18}F-PEG6-IPQA injection, PET imaging procedure, and required follow-up visits.

After enrollment, the subject will be scheduled for PET imaging using ^{18}F-PEG6-IPQA injection. Each subject in the first cohort will receive an ^{18}F-PEG6-IPQA injection with a maximum total activity of 70 Mbq. The ^{18}F- PEG6-IPQA injection will be administered by intravenous injection and followed by a saline flush.

PET imaging will comprise 3 sessions of imaging periods. The total imaging time is approximately 3 hours. Within 3 weeks prior to the PET imaging day, the subject will undergo imaging procedures as required for standard of care, e.g., CT, MRI, bone scintigraphy, X-ray, ^{18}F-FDG PET, or ultrasound. The diagnosis obtained from these examinations will be an imaging reference standard for part of the secondary efficacy evaluation.

Subjects will be included in the study for approximately 6 weeks, from signing the informed consent until 2 weeks after PET imaging visit.

The images obtained from all sessions will be used for the dosimetry and distribution analysis of the ^{18}F-PEG6-IPQA PET imaging agent. Estimates of uptake and retention in tumors will be made and compared to those in normal tissue using data from the multiple PET acquisitions. Tumors identified with ^{18}F-PEG6-IPQA whole body PET imaging will be correlated to the tumors identified with standard of care imaging examinations. Comparisons will be made on an overall (i.e., all tumor lesions) and tumor type basis. Only tumors (or metastases) identified with standard of care imaging will be used for the secondary endpoint efficacy/delectability analysis.

Safety will be assessed from the rates of adverse events (AEs), changes in vital signs, changes in ECG parameters, and changes in physical examination findings. Safety assessments will be performed at various pre- and post-treatment time points (see Table 1).

Table 1. Study Schedule of Events

	Baseline	- 72 hours	Imaging day prior to dosing	Dose	+1,3,8, and 16 min	+ 30 min	+45, 60, 90, &150 min	+ 3 hour or end of imaging	+ 24 hours	+ 1 week	+ 2 weeks
Informed consent	●										
Study entry criteria	●										
Demographic information	●										
Medical history	●										

Prior/concomitant medication	●									
[1]Physical examination	●									●
Injection site monitoring			●		●		●	●	●	●
[2]Pregnancy test females of childbearing potential)		●								
Vital signs			●		●		●	●	●	●
[3]Standard of care Diagnostic Imaging	●									
[4]Electrocardiogram	●		●				●			●
[5]Blood samples (serum biochemistry, haematology)	●							●	●	●
Adverse events (post –treatment)				●			●	●	●	●
^{18}F-PEG6-IPQA injection				●						
Blood sample acquisition					●	●	●			
[6]Urine sample acquisition										
[7] ^{18}F-PEG6-IPQA PET imaging										

[1] Physical exam includes vital signs, exam of heart, lungs, mental status, motor strength, sensory perception, pertinent organ systems and anatomic sites as medically necessary.

[2] Subjects not surgically sterile by tubal ligation or hysterectomy or amenorrheic for less than 12 months will be considered "of childbearing potential".

[3] Standard of care imaging for all tumor types must be within 21 days of protocol PET imaging

[4] Lead II ECG only will be obtained at -5 min pre-dosing and 2 hr 30 min. post-dosing. Regular 12 lead ECG will be obtained at baseline and 30 days post-dosing.

[5] Serum biochemistry: BUN, creatinine, total bilirubin, AST and ALT, alkaline phosphatase; albumin, total protein, and serum glucose. Hematology will be CBC.

[6] Urine sample acquisition: Urine will be collected when the patient has to void or at the end of imaging. Shaded area indicates continuous assessment.

[7] Imaging timeline and manual for the first cohort (short/medium height [<6 ft]) is in Appendices III and IV. 3 imaging sessions per study are performed. Each imaging session is composed of two PET/CT scans for a total of 6 scans in total per the Timelines.

Chemistry, Manufacturing, and Controls (CMC)

In collaboration with Cyclotope, we have successfully converted the production of 18F-PEG6-IPQA from a lab scale manual process to an automated cGMP process using GE Tracerlab with product-specific hardware modifications. We are in the process of making the final modifications for optimal yield. The final process will be validated thereafter. All QC methods for product release have been developed. They will be validated together with the production validation. The validated manufacturing procedure will be used to make all clinical supply in phase I study.

Key Research Accomplishments

- Completed an IND-driven GLP toxicology study of ^{18}F-PEG6-IPQA in rats. This is one of the most critical studies required by FDA before any clinical trial can be conducted using this imaging agent. The results of the study enable us to move the agent into clinical trial.
- Completed Phase I clinical protocol development. This protocol has been reviewed and approved by the Department of Nuclear Medicine where trial will be conducted, and is now pending IRB approval.
- Developed an automated manufacturing process that complies with cGMP based on the lab manual process. With this cGMP production, we are able to produce FDA-acceptable imaging agent to be used in the clinical phase I trial.
- Wrote the Investigational Drug Application (IND) to be submitted to FDA for its permission to conduct the phase I trial.

- Two manuscripts (on studies in mice and non-human primates) are in the final stages of preparation for publication.

Conclusions

Development of ^{18}F-PEG6-IPQA i nto c linical tr ial has been l argely c ompleted. We w ill nex t pursue IRB and FDA approval of the protocol and IND, respectively, before initiating the clinical phase I trial, which is anticipated during the next unfunded period (no-cost extension).

Project 3: Targeted Peptide-based Systemic Delivery of Therapeutic and Imaging Agents to Lung Cancer

(PI and co-PI: Renata Pasqualini, Ph.D., Wadih Arap, M.D., Ph.D.)

The studies outlined in this proposal focus on the use of peptide sequences with selective lung tumor-targeting properties. We will seek to validate these probes as delivery vehicles in drug and gene -targeting appr oaches. Thi s appr oach di rectly s elects *in vivo* for c irculating pr obes capable of p referential homing into tumors. The s trategy will be to c ombine homing peptides in the c ontext of phage a s gene ther apy v ectors. G iven that m any of o ur pepti des al so tar get angiogenic vasculature i n addi tion to tum or c ells, the se studies ar e l ikely to e nhance the effectiveness of therapeutic apoptosis induction and imaging technology.

Aim 1   To select peptides targeting primary and metastatic tumors in lung cancer patients.

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 2   To validate receptors for targeting human lung cancer.

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 3   To design tools for molecular imaging of lung tumors.

Summary of Research Findings
In collaboration with Dr. Juri Gelovani, our group has previously reported the design, generation, and construction of AAV/phage (termed AAVP) particles (Hajitou et al. 2006, Hajitou et al. 2007, Soghomonyan et al . 2 007) for targeted m olecular-genetic i maging. Thes e hy brid v ectors containing prokaryotic and euk aryotic c is-genomic el ements hav e the potenti al to i ntegrate ligand-directed targeting and molecular-genetic imaging. In a related line of research, we have used labeled targeted peptide motifs themselves as imaging tools (Yao et al . 2005, M archiò et al. 2004, A rap et al . 2004, Zurita et al. 2004, C ardó-Vila et al. 2003, C hen et al . 2003, Mintz et al. 2003). In pilot experiments, AAVP-based molecular-genetic imaging appears to be superior to FDG i n side-by-side c omparisons because it provides prediction of therapeutic response in addition to response monitoring (Hajitou et al., *PNAS*, 2008). Thus, we plan focus primarily on the dev elopment of A AVP-based m olecular-genetic i maging. Fi nally, w e hav e al so des igned and developed nanotechnology-based (i.e., bottom-up self-assembled) biocompatible networks of phage -gold as nanot echnology-based molecular sensors and r eporters (Souza et al . 2006a,

Souza et al. 2006b); this new methodology will be incorporated and it will likely prove to be quite synergistic with AAVP (Souza et al. Nature Nanotechnology, 2010).

Future Planned Research
There are several areas of research planned: (i) To use prototypes of this new class of targeted hybrid vectors for therapy and for molecular-genetic imaging; (ii) to develop AAVP-based library applications; (iii) to create other chimeric prokaryotic-eukaryotic vectors; and ultimately (iv) to generate an "imaging transcriptome" for lung cancer.

(i) We will use targeted prototypes of this class of hybrid vectors for molecular-genetic imaging/therapy, specifically for:
- Discovery of new ligand motifs that target human tumor endothelium (Staquicini et al., in progress)
- AAVP-based anti-vascular cancer therapy by targeted TNF in pet dogs with native tumors (Paoloni et al., in press)
- Integrate biocompatible networks to create a transducing matrix (Driessen et al., in preparation)

(ii) To develop AAVP-based combinatorial peptide libraries for use in directing patient selection, e.g., for patient settings such as pre-operative and with metastatic tumors. As such, the steps towards this goal are as follows:
- Design and production of targeted AAVP prototypes and libraries in GMP-facilities for patient applications.
- NIH RAC approval for long-term transduction in cancer patients.
- Proof-of-concept with a reporter/suicide gene (i.e., HSV-*tk*) or targeted TNF.

(iii) We will create other hybrid vectors with the biologic attributes of bacteriophage and animal viruses. Generation of a double-stranded DNA construct with elements of adenovirus and of lambda phage is ongoing (Sun et al., in progress)

(iv) The incorporation of transcriptional targeting (through tissue-specific or radiation-induced promoters) to ligand-directed AAVP-targeting may enable one to determine a gene's (or set of genes') status without tissue biopsy.

Mentoring, Past Outcomes, and Future Training
Several investigators have developed either into academic diagnostic radiologists (e.g., Bradley Restel, M.D.) or became independent laboratory-based Principal Investigators with an interest in molecular imaging (Amin Hajitou, Ph.D., Glauco Souza, Ph.D.). The current and next-generation of investigators in training include Suren Soghomonyan, Ph.D. and Michael Ozawa, M.D./Ph.D. student. We have requested a cross-appointment within the Department of Experimental Molecular Imaging (with Dr. Gelovani). We believe that a formal appointment within this department will mutually enhance the productivity of a Targeted Imaging Program (TIP) among other translational research initiatives. We also trust that reportable outcomes including improved cross-training of translational researchers in molecular imaging will be facilitated as a consequence of this faculty appointment, and so a series of meetings and standing related activities has been planned for mid- to late-2010.

Key Research Accomplishments

- Used labeled targeted peptide motifs themselves as imaging tools.
- Demonstrated that AAVP-based molecular-genetic imaging appears to be superior to FDG in side-by-side comparisons for predicting therapeutic response.

- Designed and dev eloped nanote chnology-based (i.e., bottom -up, self-assembled) biocompatible networks of phage-gold as nanotechnology-based molecular sensors and reporters.

Conclusions

The c entral w orking hy pothesis in our pr ogram i s that di fferential pr otein ex pression i n the human vascular endothel ium as sociated w ith l ung c ancer offer s the potential for dev eloping novel diagnostic, imaging, and therapeutic s trategies. In es sence, c ombinatorial l ibrary selections (peptide- and antibody-based) are leveraged to discover, validate, an d ex ploit th e vascular bi ochemical di versity of endothelial cell s urfaces tow ards a new v ascular-targeted pharmacology. Such targeting technologies may lead to the development of ligand-directed agents for appl ication i n the tr eatment of c ancer pati ents. Tr anslational appl ications, s uch as first-in-human clinical trials, have now began within the institution, as the Food and Drug Administration (FD A) has r ecently gr anted a "safe-to-proceed" s tatus for the fi rst v ascular-targeted Inv estigational New D rug, discovered, dev eloped and bei ng e valuated i n pati ents at MDACC. Such trials will ultimately determine the value of this strategy. Two other drugs are in pre-IND stage and several others in pre-clinical laboratory phase. Long-term, the broader vision of the r esearch i s a l arge-scale m apping of r eceptors i n human v asculature to wards a new ligand-directed pharmacology.

Project 4: Inhibition of bFGF Signaling for Lung Cancer Therapy

(PI: Reuben Lotan, Ph.D.)

The survival of lung cancer patients is poor because this cancer is diagnosed at advanced stages. Th erefore, i mprovements i n ear ly d etection thr ough the i dentification of m olecular markers for diagnosis and for intervention combined with targeted chemoprevention are urgently needed. While the m olecular events involved in lung cancer pathogenesis are being unraveled by ongoi ng l arge s cale genom ics, proteomics, and m etabolomics s tudies, i t i s a lready w ell recognized that proliferation-, s urvival- and angi ogenesis- promoting s ignaling pa thways ar e amplified i n l ung cancer. A mong the angi ogenesis s ignaling pathw ays, the bas ic fi broblast growth factor (bFGF) and its transmembrane tyrosine kinase receptors (FGFRs) are playing important roles in addition to the well-studied vascular endothelial growth factor (VEGF) and its receptors (VEGFRs). B oth ty pes of angi ogenesis s ignaling pathways, the V EGF/VEGFR an d the bFGF/FGFR, have been detected in NSCLC and associated with lung cancer development. However, most efforts in preclinical and c linical trials have been di rected to the V EGF/VEGFR pathway.

We hypothesize that bFGF triggers signaling pathways that contribute to malignant progression of l ung c ancers by s timulating tum or c ell and endothel ial c ell pr oliferation and s urvival an d augmenting angi ogenesis. Ther efore, agents that i ntervene in thi s pathway may b e useful for lung c ancer ther apy ei ther al one or i n c ombination w ith agents that t arget the V EGF/VEGFR signaling pathways and/or with cytotoxic agents. We will address the following specific aims in order to under stand the mechanism(s) underlying the *in vitro* and *in vivo* effects of bFGF on l ung cancer and endothelial cells and the ability of bFGF inhibitors to suppress the growth of NSCLC *in vitro* and *in vivo.*

Aim 1   Determine the effects of bFGF on *in vitro* growth, survival, motility, invasion and angiogenesis of NSCLC cells and endothelial cells.

Summary of Research Findings
This aim was completed and summarized in the previous reports.

Aim 2   **Evaluate the relative potency of several inhibitors of bFGF binding to receptor (i.e., TMPP and analogs) in inhibiting effects of bFGF detected in Specific Aim 1 and evaluate the effects of these inhibitors in combination with paclitaxel on *in vitro* growth and survival of tumor cells.**

Summary of Research Findings
This aim was completed in conjunction with Aim 4. Results are reported under Aim 4.

Aim 3   **Evaluate anti-tumor activity (growth inhibition, apoptosis, suppression of angiogenesis) of the most effective inhibitor identified in Specific Aim 2 when used alone and in combination with paclitaxel in an orthotopic lung cancer model using luciferase-expressing NSCLC cells for *in vivo* bioluminescence imaging of tumor growth and response to treatment.**

Summary of Research Findings
As previously reported, we were unable to complete this specific aim because development of the FGF signaling inhibitory small synthetic molecules was discontinued (e.g., TMPP by Prochon Biotech and SSR128129 by Sanofi/Aventis), and other companies with such inhibitors (e.g., FGFR2-FC from Centilion co and BMS-582664 from Bristol Myer Squibb) were unwilling to provide sufficient quantities of their compounds for *in vitro* and animal studies.

Aim 4   **To investigate the expression of bFGF signaling components (bFGF, FGFR-1, FGFR-2, heparan sulfate, syndecan-1, and FGFR-3) by IHC staining of tissue microarrays (TMAs), and correlate the expression of bFGF/bFGFRs between tumor and non-malignant epithelial cells with angiogenesis.**

Summary of Research Findings
Over the previous year, our research has focused on understanding the mechanism of the anti-tumor effects of agents that target FGFR1-mediated signaling, especially the adenoviral vector expressing the dominant-negative FGFR1 construct (AdV/DNFR1). We focused on this agent after demonstrating, as reported previously in last year's report, that it had a potent growth inhibitory effect on several lung cancer cell lines. We explored the potential role of the heat shock protein Hsp90 as downstream effectors of FGF signaling.

Differential growth inhibitory effects of adenoviral DNFR on normal and malignant lung cancer cells
We compared normal human bronchial epithelial cells (NHBE) to those of the malignant lung cell lines (1170-I, A549, and H1299) regarding their response to inhibition of FGFR1 signaling using an adenoviral vector containing the dominant-negative FGFR1 (AdV/DNFR1). We found that the malignant cells were more sensitive to growth inhibition than the normal cells (Table 2). NHBE and malignant lung cancer cells were treated with adenoviral vector containing only LacZ reporter or the dominant negative

| Table 2. Effects of AdV/DNFR1 on cell growth		
Cells	Growth inhibition by AdV/DNFR1	
	48 hours	72 hours
Normal NHBE	10.0	16.2
1170-I	23.1	32.5
A549	21.3	42.0
H1299	20.8	39.8

FGFR1 construct or without any treatment, and the cell numbers were determined 48 and 72 hours later using the colorimetric sulforhodamine B assay. Growth inhibition as a percentage of the untreated cells was calculated and the inhibition by the LacZ vector was subtracted from that of the AdV/DNFR1.

Table 3. Effects of AdV/DNFR1 on cell cycle and apoptosis in H1299 cells		
Cell cycle phase	Control	AdV/DNFR1
G1	38.7	26.0
S	38.1	20.5
G2/M	23.2	53.5
Apoptosis	3.22	14.7

The growth inhibition was primarily the result of induction of a cell cycle arrest in the G2/M phase and, to a much lesser degree, via induction of apoptosis (Table 3). The malignant lung cancer cells H1299 were treated with adenoviral AdV/DNFR1 or control vector and, after 48 hours, they were harvested and their DNA was stained with propidium iodide or with the apoptosis TUNEL reagents; cell cycle and apoptosis were then analyzed by flow cytometry.

Mechanisms of cell cycle arrest in G2/M by AdV/DNFR1 infection

To understand the mechanism by which suppression of FGFR1 signaling using the AdV/DNFR1 inhibits growth by G2/M cell cycle block, we analyzed various proteins that regulated cell cycle in 1170-I cells treated with AdV/GFP control vector or AdVDNFR1. Cdc2, a cyclin- dependent kinase (also called CDK1), determines the onset of mitosis in all eukaryotic cells. In its unphosphorylated state, CDK1 forms a complex with cyclin B1 to trigger entry into mitosis. The function of cdc2 is inhibited by phosphorylation on tyrosine 15, which causes dissociation from cyclin B1. This phosphoryation can be caused by the kinase Wee1 while the phosphatase cdc25 can dephosphorylate cdc2. Thus, an increase in Wee1 and decrease in cdc25 can each induce a G2
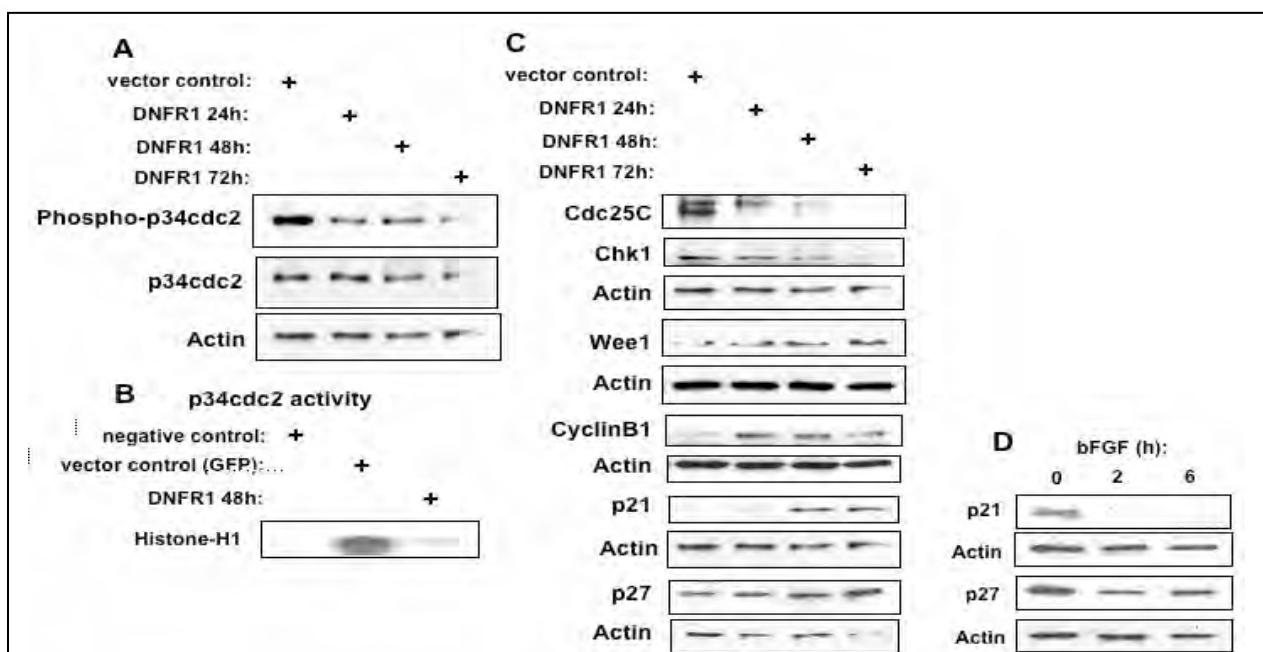

Figure 9. *A*: Analysis of cell cycle related proteins in AdV / DNFR1 infected 1170-I cells. The cells were infected with vector control (GFP) or the dominant negative vector and after 24, 48 and 72 hours, the cells were harvested and lysates were prepared. ***A and C***: The lysates were separated by 10% SDS-PAGE and analyzed for levels of the indicated proteins by Western blotting. ***B***: The lysates were subjected to *in vitro* kinase assay by immunoprecipitating p34cdc2 followed by analysis of its kinase activity using histone H1 as a substrate and of [^{32}P] adenosine triphosphate (ATP) as phosphate donor followed by gel electrophoresis and autoradiography of the dried gel. ***D:*** The 1170-I cells were treated with vehicle or 20 ng/ml bFGF for 2 or 6 hours in KGM medium (without supplement). The cells were lysed and analyzed by Western blotting for levels of p21, p27 and actin.

arrest. Chk1 kinase can activate Wee1 and inactivate Cdc25 phosphatase and these effects can lead to G2 arrest as well. P21 can inhibit cdc2-cyclin B complex and block cell cycle in G2. As shown in Figure 1, we found that AdVDNFR1 decreased the expression of p34cdc2 mainly after 48 and 72 hours, but more rapidly decreased the level of cdc2 phosphorylation, after only 24 hr (Figure 9A). Furthermore, the kinase activity of cdc2 was suppressed in the DNFR1 infected cells (Figure 9B). In addition, the expression of DNFR1 decreased the levels of cdc25 and Chk1 and increased Wee1, and the cyclin–dependent kinase inhibitors (CKI) p21 (inhibits cdk2 and cdk4) and p27 (CDKN1B) (Figure 9C). These changes can explain the G2 arrest. The increase in cyclin B1 is not clear but, with a low level of cdc2, cyclin B1 would not be able to function in triggering mitosis. Interestingly, the effect of the DNFR1 was opposite to the effect of bFGF, which suppressed the levels of the cell cycle kinase inhibitors (CKI) p21 and p27 (Figure 9D). The opposite effects of DNFR1 and bFGF lend further support to the conclusion that DNFR1 blocks the FGFR1 signaling. These results indicated that the adenoviral vector containing a dominant negative FGFR1 receptor construct could potentially be useful for treatment of lung cancer *in vivo*. Studies targeting lung cancer could benefit from an aerosolized delivery by inhalation.

Induction of Hsp90 by bFGF and effects of Hsp90 inhibition on lung cancer cell growth
Other reported that the mitogenic effects of bFGF in breast cancer cells require the presence of the heat shock protein Hsp90. Hsp90 is ubiquitous but its expression is elevated in many cancer cells, including lung cancer compared to normal cells. It regulates cellular stress responses by acting as a chaperon protein; namely, it preserves the function of proteins by maintaining their conformation and stability. These client proteins play important functions such as growth, differentiation, and survival. Initially, we determined whether Hsp90 and other Hsps are expressed in lung normal and malignant cells. We found that only Hsp90 showed a consistent increased level between normal and malignant cells, and its level increased gradually from normal through immortalized and transformed cells to malignant cells (Figure 10).
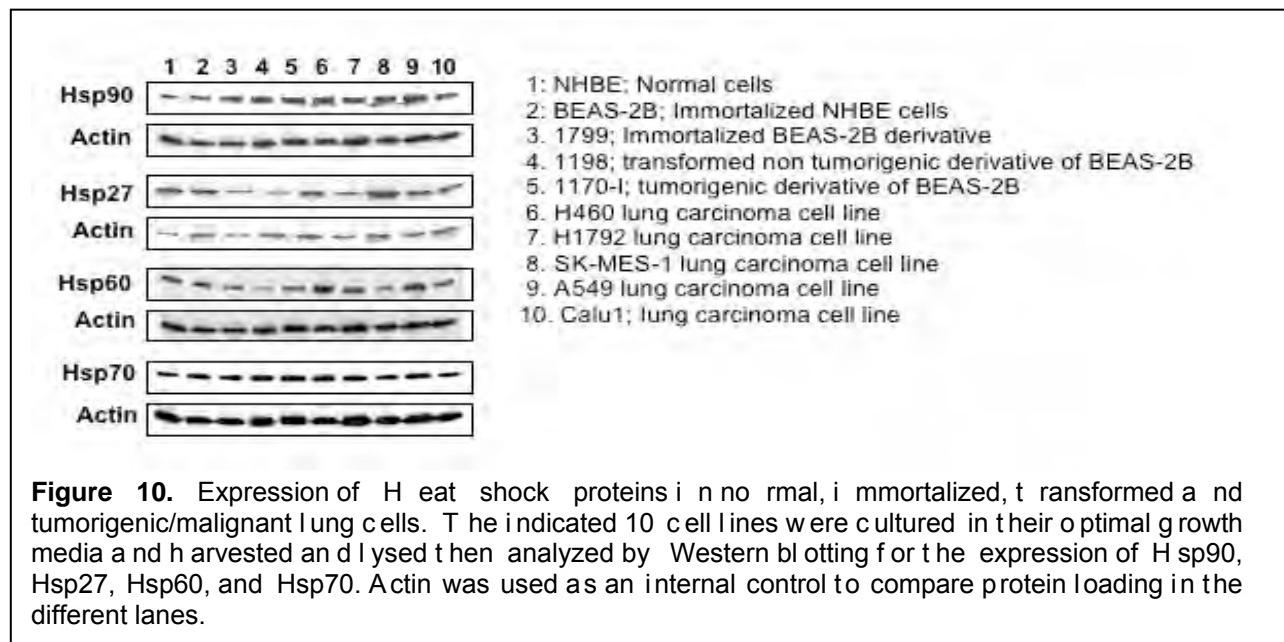
Figure 10. Expression of Heat shock proteins in normal, immortalized, transformed and tumorigenic/malignant lung cells. The indicated 10 cell lines were cultured in their optimal growth media and harvested and lysed then analyzed by Western blotting for the expression of Hsp90, Hsp27, Hsp60, and Hsp70. Actin was used as an internal control to compare protein loading in the different lanes.

To determine whether Hsp90 is important for the growth of lung cancer cells and whether its expression is related to the effects of bFGF and to FGFR signaling in lung cancer cells, we analyzed the effects of the Hsp90 inhibitor Geldanamycin on the lung cancer cells in the absence and presence of bFGF. We found that treatment of lung cancer cells with Geldanamycin inhibited

cells growth by blocking cells in the G2 phase of the cell cycle and by inducing apoptosis (Figure 11A). Treatment with a low dose of 50 nM Geldanamycin was able to decrease the percentage of cells i n the S phas e and doubl e the c ells i n th e G 2/M ph ase. H owever, the ov erall effec t o n apoptosis was rather small (<10%). Treatment with a high dose of 1 µM Geldanamycin induced a more substantial decrease in S phase, an increase in G2/M (Figure 11A, lower panels), and a particularly impressive increase in apoptosis up to 76.8% (Figure 11B, lower panels).
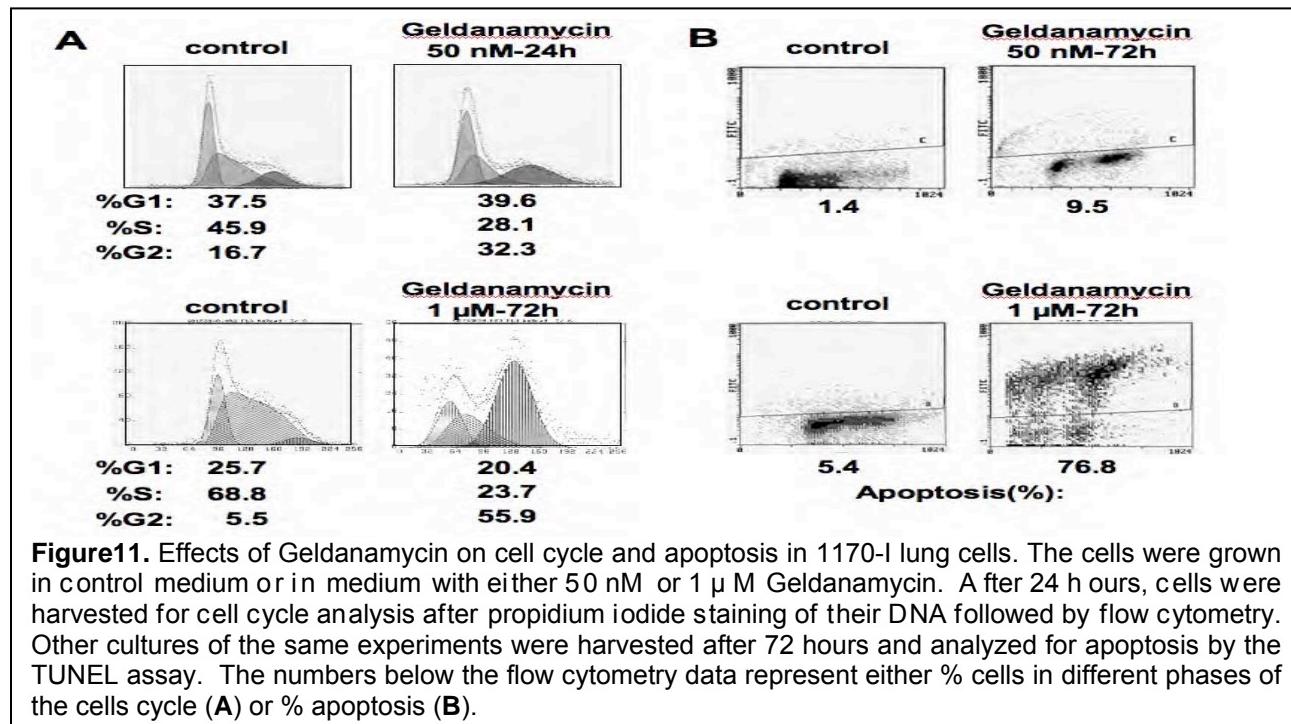
Figure11. Effects of Geldanamycin on cell cycle and apoptosis in 1170-I lung cells. The cells were grown in control medium o r i n medium with either 5 0 nM or 1 µ M Geldanamycin. A fter 24 h ours, c ells w ere harvested for cell cycle analysis after propidium iodide staining of their DNA followed by flow cytometry. Other cultures of the same experiments were harvested after 72 hours and analyzed for apoptosis by the TUNEL assay. The numbers below the flow cytometry data represent either % cells in different phases of the cells cycle (**A**) or % apoptosis (**B**).

To explore the mechanism of the effects of Geldanamycin, we analyzed the level of proteins related to cell cycle regulation and apoptosis and found that Geldanamaycin increased the level of cyclin B 1 a nd p21 and dec reased the l evel o f Wee1 (Figure 12). I mportantly, G eldanamycin decreased the level of cdc2 (Figure 12B) and suppressed the kinase activity (Figure 12C). The latter effects were reminiscent of the effects of the dominant-negative FGFR1 infection (Figure 9).
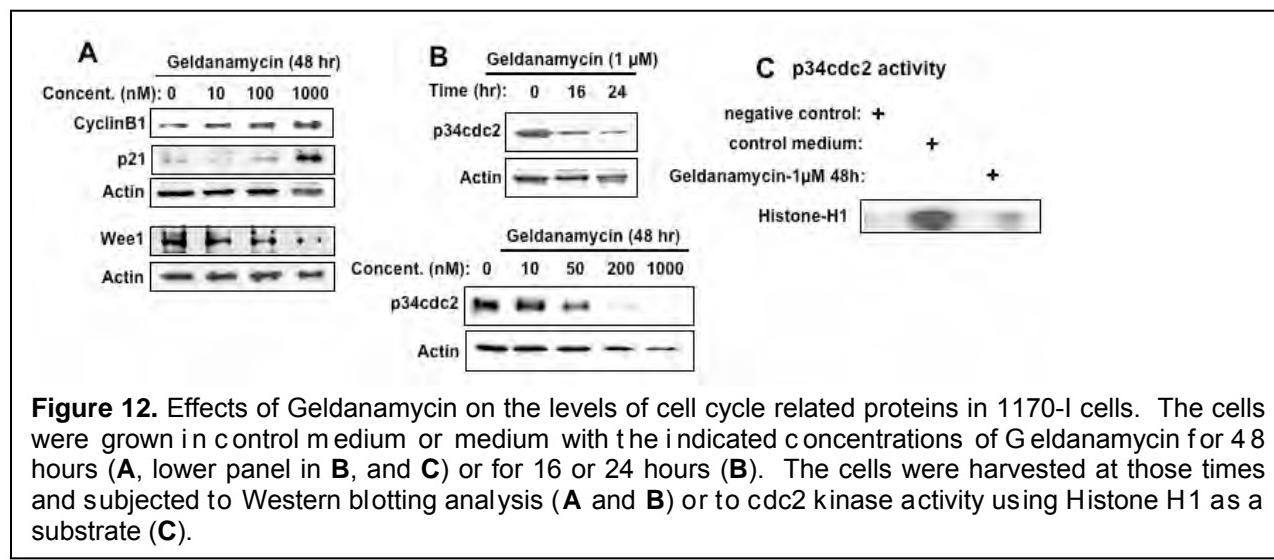

Figure 12. Effects of Geldanamycin on the levels of cell cycle related proteins in 1170-I cells. The cells were grown i n c ontrol m edium or medium with t he i ndicated c oncentrations of G eldanamycin f or 4 8 hours (**A**, lower panel in **B**, and **C**) or for 16 or 24 hours (**B**). The cells were harvested at those times and subjected to Western blotting analysis (**A** and **B**) or to cdc2 kinase activity using Histone H1 as a substrate (**C**).

To explore the possible cross-talk between Hsp90 and FGF signaling, we analyzed the effect of bFGF treatment on Hsp90 level and found that treatment of lung cancer cells with bFGF induces the heat shock protein Hsp90 without affecting the levels of other Hsp protein like Hsp27 and Hsp70 (Figure 13).

Figure 13. Induction of heat shock protein Hsp90 by bFGF in 1170-I cells. The cells were treated for 2 hours with the indicated concentrations of bFGF then harvested and analyzed by western blotting for the expression of Hsp27, 70 and 90. The same membrane was probed with actin as a control for protein loading.
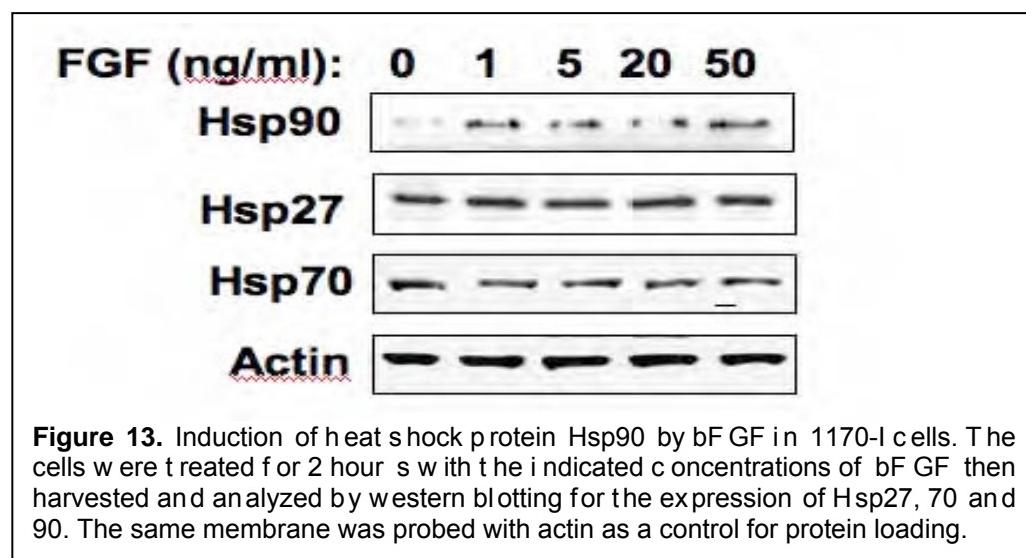
To assess the possible role of Hsp90 in the growth stimulatory effects of bFGF on the lung cells, we examined the ability of Geldanamycin to affect the growth stimulation of the lung cells by bFGF. We found that Geldanamycin inhibited, in a dose-dependent fashion, the growth stimulatory effect of bFGF in 1170-I cells (Figure 14). These results suggest that Hsp90 function is important for the mitogenic effects of bFGF in the lung cells.

Figure 14. Effects of Geldanamycin on the mutagenic effect of bFGF. 1170-I cells were grown in control medium or medium with 20 ng/ml bFGF without and with the indicated concentrations of Geldanamycin for 72 hours (**A**) or 18 hr (**B**). The cells were harvested at those times and subjected to cell counting by a colorimetric assay (**A**) or to cell cycle analysis by propidium iodide DNA staining followed by flow cytometry (**B**).

Key Research Accomplishments

- Discovered that blocking the FGFR1 signaling by DNFR1 leads to inhibition of cell growth by blocking cells at the G2 phase of the cell cycle through decreased expression, phosphorylation, and kinase function of p34cdc2 combined with increased Wee1, and the cyclin-dependent kinase inhibitors p21 and p27.
- Determined that Geldanamycin inhibits the growth of lung cancer cells by a G2/M block and by a mechanism that in part emulates the blocking of FGFR signaling by DNFR1 expression.
- Demonstrated the ability of bFGF to increase Hsp90 level and the ability of the Hsp90 inhibitor Geldanamycin to inhibit FGF signaling, suggesting that Hsp90 is a positive mediator of FGF effects.

Conclusions

Our results indicate that the adenoviral vector containing a dominant negative FGFR1 receptor construct, which inhibits FGFR signaling, could potentially be useful for treatment of lung cancer *in vivo.* Studies targeting lung cancer with this adenoviral construct are warranted. Such studies could focus on adenovirus delivery by inhalation using an aerosolized preparation. In terms of understanding FGFR signaling, our studies highlighted an important role for Hsp90 as a downstream mediator of FGF effects. This suggests that combined targeting of FGFR signaling (e.g., by AdVDNFR1) and Hsp90 function (e.g., by Geldanamycin or other Hsp90 inhibitors) could provide additive or synergistic efficacy.

Project 5: Targeting mTOR and Ras signaling pathways for lung cancer therapy

(Project Co-leaders: Shi-Yong Sun, Ph.D., Suresh Ramalingam, M.D.)

It should be noted that Dr. Suresh Ramalingam assumed the leadership of Project 5 following Dr. Fadlo Khuri's decision to step down due to increased administrative responsibilities; this administrative change was approved by the DoD in November 2009. Following is a summary of our research progress:

Aim 1   To determine whether an mTOR inhibitor inhibits the growth of human NSCLC cells via G1 growth arrest or induction of apoptosis, and to identify the molecular determinants of mTOR inhibitor sensitivity.

Summary of Research Findings
This aim was completed and summarized in the previous reports.

Aim 2   To determine whether the effect of mTOR inhibitors on the growth of human NSCLC cells is enhanced in the presence of a PI3K inhibitor or a MAPK inhibitor.

Summary of Research Findings
The initial proposed studies have been completed and summarized in the previous reports. Additionally, we determined whether or not the sequence of the combination treatment affects the therapeutic outcome. We found that the concurrent treatment with rapamycin and LY294002

was more effective than sequential treatments with either rapamcyin followed with LY294002 or LY294002 followed by rapamcyin in inhibiting the growth of cancer cells (Figure 15).
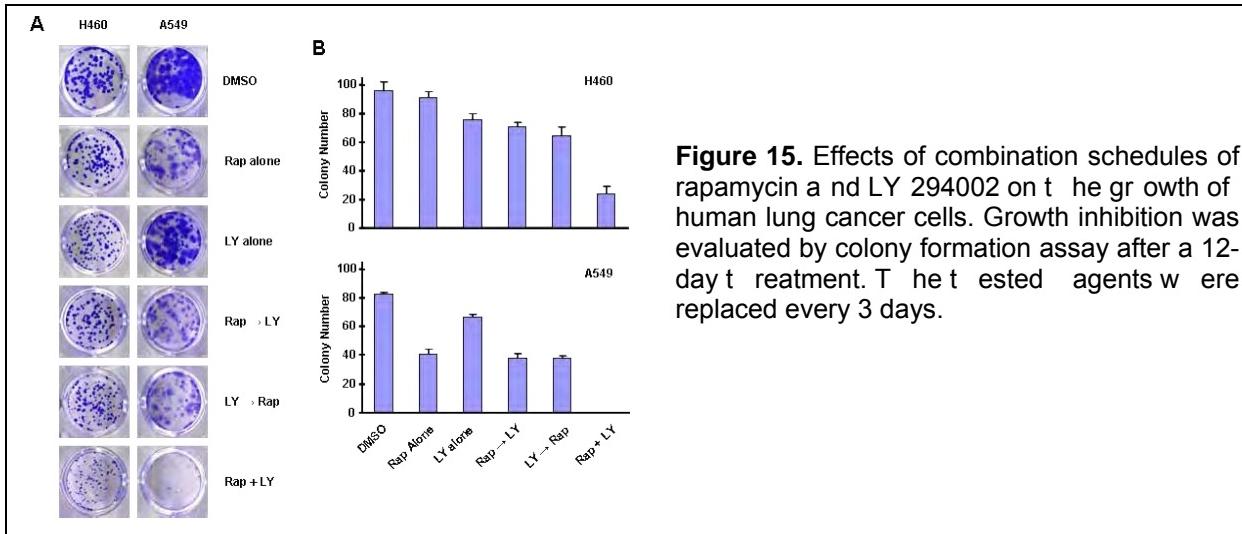
Figure 15. Effects of combination schedules of rapamycin a nd LY 294002 on t he gr owth of human lung cancer cells. Growth inhibition was evaluated by colony formation assay after a 12-day t reatment. T he t ested   agents w ere replaced every 3 days.

We obtained the novel PI3K/mTOR dual inhibitor BEZ235, and we further examined the effects of RAD001 combined with BEZ235 on the growth of human lung cancer cells. We found that the combination of sub-optimal concentrations of RAD001 and BEZ235 exhibited synergistic effects on inhibiting the growth of the lung cancer cells evaluated. In agreement with the rapamycin and LY294002 c ombination, concurrent combination of R AD001 and BEZ235 was m ore effec tive than s equential combination tr eatment i n i nhibiting the growth of c ancer c ells (Figure 16). Interestingly, we found that p-Akt levels were not reduced in the cells treated with the combination, s uggesting that the   combination of R AD001 and B EZ235 ex hibits enhanc ed growth inhibitory effects of cancer cells without inhibition of Akt. The ongoing work will focus on elucidating the underlying mechanism.
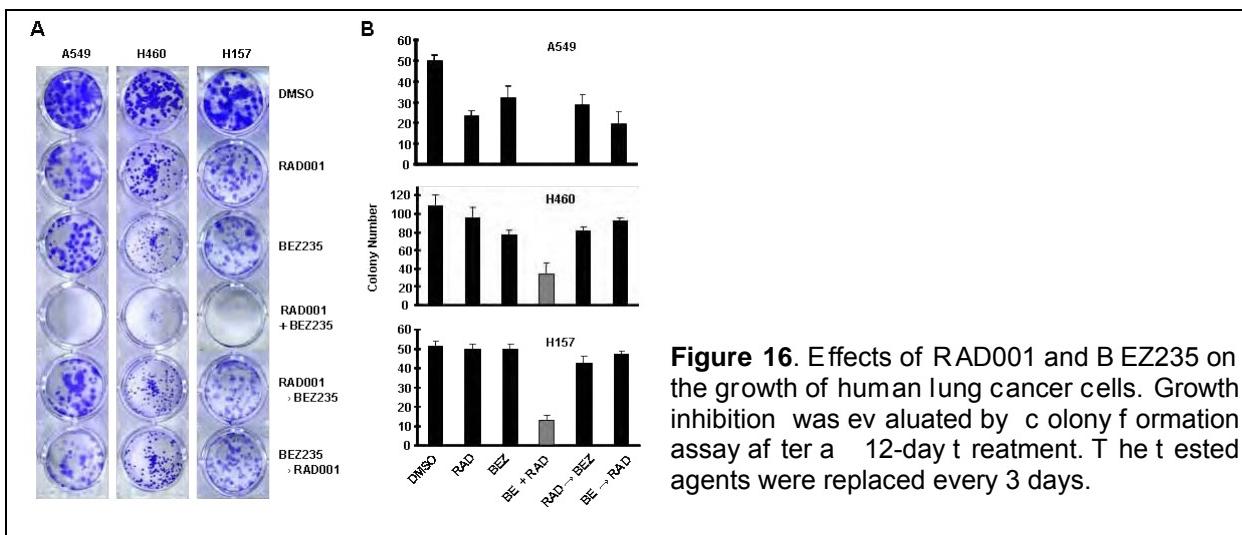
Figure 16. E ffects of R AD001 and B EZ235 on the growth of hum an l ung c ancer c ells. Growth inhibition was ev aluated by c olony f ormation assay af ter a   12-day t reatment. T he t ested agents were replaced every 3 days.

Aim 3   To evaluate the efficacies of the combinations of rapamycin with LY294002 or U0126 in nude mice models of lung cancer xenografts *in vivo*.

Summary of Research Findings

This aim was completed as reported in the previous annual report.

Aim 4   To conduct a pilot clinical biochemical induction trial to investigate the effect of RAD001 in operable NSCLC patients and identify molecular determinants of RAD001 sensitivity and prognosis.

Summary of Research Findings

A phase IB study to evaluate the molecular effects of everolimus (RAD001) on non-small cell lung c ancer tum ors i s c urrently under way. P atients w ith surgically r esectable NSCLC ar e eligible for t he s tudy. A fter an i nitial tum or bi opsy and a P ET s can, p atients ar e t reated w ith everolimus for 3 w eeks at a dos e of 5 m g P O Q D. O n day 22, pat ients under go s urgical resection. T he tum or ti ssue at ba seline and   the pos t-therapy s pecimens ar e anal yzed fo r various biomarkers in the AKT-mTOR pathway. A total of 11 patients have been enrolled to the study including one patient on the control arm. The median age of the pa tients is 60 years. Five patients had adenocarcinoma histology and one had squamous cell carcinoma. The tr eatment was tolerated well overall for the 10 patients on the tr eatment arm. The salient adverse events included grade 1/2 nausea, emesis, and hypertriglyceridemia. Two patients had prolonged post-operative recovery that was attributed to pre-existing co-morbid illness.
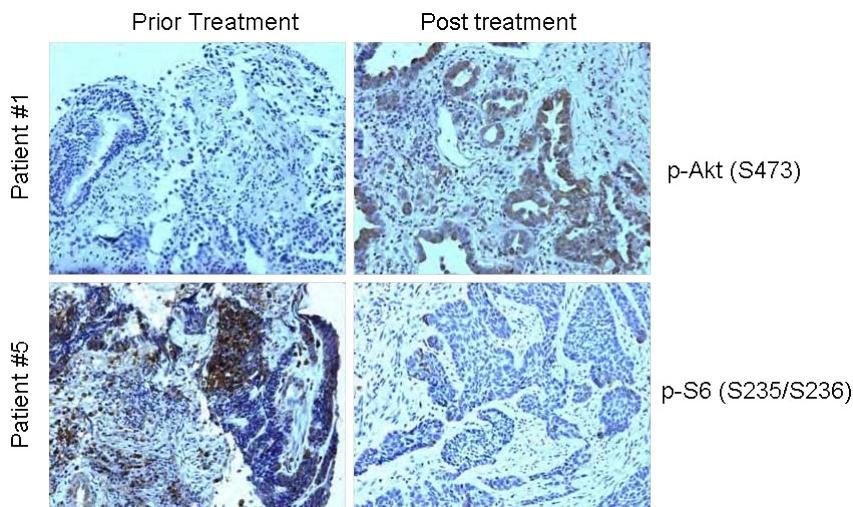
Figure 17. Staining of p-Akt in patient #1 and p-S6 in patient #5 using immunohistochemistry.

Five out of 6 patients, with available paired PET scans, demonstrated a reduction of 15 - 42% in standardized uptake v alue (SUV) for FD G l evels. Im munohistochemistry s tudies w ere performed o n the pai red tum or ti ssues for 6 pati ents. An i ncrease in p -AKT was n oted in the post-treatment tumor tissue in 5 out of 6 patients (Figure 17). This is consistent with our pre-clinical s tudies that de monstrated c onsistent upregulation of A KT w ith i nhibition of m TOR signaling (Sun et al, Cancer Res, 2005). Phosphorylated S6 was downregulated in one out of 6 specimens (Figure 17) . O nly one   out of 5 pa tients dem onstrated a r eduction i n p -P70S6k activity w ith the ev erolimus ther apy. Whi le s ome of thes e m olecular c hanges d ocument the biological effects of m TOR inhibition on the tu mor tissue, some markers were not affected and this could be r elated to the time el apsed between the l ast dose of everolimus i ngested by the patient and the ti me of s urgery. C ontinued ac crual to t he s tudy w ill hel p unde rstand thes e

recorded events; we intend to accrue a total of 20 patients to the treatment arm and 10 to the control arm.

Key Research Accomplishments

- Demonstrated that concurrent treatment with rapamycin and LY294002 is more effective than sequential treatments with either rapamcyin followed with LY294002 or LY294002 followed by rapamcyin in inhibiting the growth of cancer cells.
- Discovered that the combination of sub-optimal concentrations of RAD001 and BEZ235 exhibited synergistic effects on inhibiting the growth of the lung cancer cells evaluated.
- Enrolled a total of 11 patients to the study, including one patient on the control arm.

Conclusions

During the previous project period, we continued to evaluate whether the effect of mTOR inhibitors on the growth of human NSCLC cells is enhanced in the presence of a PI3K inhibitor or a MAPK inhibitor. We demonstrated that the combination of sub-optimal concentrations of RAD001 and BEZ235 exhibited synergistic effects on inhibiting the growth of the lung cancer cells and that the combination of RAD001 and BEZ235 was more effective than sequential combination treatment in inhibiting the growth of cancer cells. The findings confirm the feasibility of conducting 'window of opportunity' studies as a novel mechanism to understand the molecular effects on targeted agents on the tumor tissue. We continued to enroll patients to the RAD001 trial to investigate the effect of this drug in operable NSCLC patients and to identify molecular determinants of RAD001 sensitivity and prognosis.

Project 6: Identification and Evaluation of Molecular Markers in Non-Small Cell Lung Cancer (NSCLC)

(PI and co-PI: Ralf Krahe, Ph.D., Li Mao, M.D)

A better understanding of the lung cancer biology and an identification of genes involved in tumor initiation, progression and metastasis are an important first step leading to the development of new prognostic markers and targets for therapy. In the same context, identification of reliable predictive markers for response or resistance to therapy in NSCLC patients is also desperately desired for optimal delivery of targeted therapy and/or standard chemotherapy. The proposed studies aim to identify the two types of markers that would eventually help develop smarter clinical trials, which will selectively recruit patients who are more likely to respond to one regimen over another and lead to improvement of overall therapeutic outcomes.

Aim 1   To expression profile by DNA microarray technology aerodigestive cancers - with primary focus on adenocarcinoma and squamous cell carcinoma (SCC) of the lung, and head and neck squamous cell carcinoma (HNSCC), including primary tumors and normal adjacent tissue, and (where available) metastatic lesions.

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 2   To DNA profile the same samples by complementing DNA approaches to stratify RNA expression profiles on the basis of their corresponding DNA profiles.

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 3   To evaluate the contribution of promoter hypermethylation and transcriptional inactivation of known cancer genes subject to epigenetic silencing to cancer phenotype.

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 4   To determine protein signatures of treatments of erlotinib and other therapeutic agents, alone or in combination, in NSCLC and identify molecular predictors of response.

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 5   To determine a clinical utility of the molecular predictors.

Summary of Research Findings
This aim was completed and summarized in previous reports.

Conclusions
This project was completed as reported in the previous annual report.

Core B: Biostatistics & Data Management Core

(Core Director: J. Jack Lee, Ph.D.)

The Biostatistics and Data Management Core has continued to work with all IMPACT Projects in their r esearch effor ts, especially i n the ar ea of bi ostatistical s upport i n c linical trial des ign, implementation, and analysis of experimental results. We also developed statistical methods to enhance the design and analysis pertinent to the lung cancer research.

Specific Aims:
1. To ensure that the r esults of al l pr ojects ar e ba sed on w ell-designed ex periments and are appr opriately i nterpreted by pr oviding ex perimental des ign; s ample s ize es timates; power calculations; and integrated, comprehensive analysis for each basic science, pre-clinical, and clinical study.
2. To dev elop a data m anagement s ystem that i ntegrates clinical, pathological, and b asic science data while providing data integrity through process tracking and quality control.
3. To pr ovide statistical a nd data m anagement s upport for genom ic and i maging s tudies including microarray, proteomics, and molecular targeted imaging.

4. To develop and adapt innovative statistical methods pertinent to biomarker-integrated translational lung cancer studies.
5. To produce statistical reports for all projects.
6. To collaborate and assist all project investigators with the publication of scientific results.

Summary of Research Findings

Core B continues to provide statistical support in the conduct and interim efficacy and toxicity analyses of the clinical trial proposed in Project 1: "A Phase I/II Study of Tarceva (erlotinib) in Combination with Chemoradiation in Patients with Stage III A/B Non-Small Cell Lung Cancer" (PI: Dr. Ritsuko R. Komaki). In addition, we continue to provide statistical support for the clinical trial in the Developmental Research Project 1 titled: "Treatment of Malignant Pleural Effusion with ZD6474 a Novel Vascular Endothelial Growth Factor Receptor (VEGFR) and Epidermal Growth Factor Receptor (EGFR) Tyrosine Kinase Inhibitor" (PI: Dr. Roy Herbst). Both studies have continued to enroll patients over the past year and we are anticipating that these trials will be completed and ready for statistical analysis by the end of 2010 (note: a request for a no-cost extension for the IMPACT grant has been submitted to the DoD).

Core B has worked closely with Drs. Juri Gelovani, David Stewart, and Bijun Yang in the development of a three-stage design for the protocol: "A Phase I Study of 18-F-Fluoroacetate Sodium as a PET Imaging Agent for Tumor Detection" (2009-0157). Comprehensive toxicity data from primate studies (performed outside the scope of the IMPACT grant with non-DoD funds) were thoroughly analyzed to establish a safe radiation dose for 18-F-PEG6-IPQA as a PET imaging agent for tumor detection in preparation for the IND application for submission to FDA.

We also worked with Dr. Wistuba in Core C on statistical analysis of a number of IHC studies including EGFR/EGFR ligand expression arising from studies proposed in this grant program. The biomarkers analyzed include EGF, Amphiregulin, TGF-alpha, HER2, p-HER3, HER3, EGFR, pEGFR, GRP78, IL11R EphA5, TCF21, STAT1, LCK, DUSP6, EGFR, KRAS, VEGFR, and EGFR.

Key Research Accomplishments

- Continued to provide statistical support in the clinical trials for Project 1 and DRP-1.
- Provided statistical support for Projects 2, 3, 4, 6, and Pathology Core.
- Continued to work closely with the Project 4 (PI: Dr. Reuben Lotan) on synergy studies of combination drug treatment in cell lines to determine whether the effect is synergistic, additive, or antagonistic.
- Developed methods and provided codes to construct the confidence interval for the interaction index for the Emax model with Dr. Lotan in Project 4. Further enhancements of estimating the parameters of the Emax model were provided by interpolation and extrapolation to improve the model convergence for cases when the observed data points do not yield adequate model fit. Sensitivity analysis was implemented for checking the model robustness as well.
- Developed a Bayesian dose-finding trial design with multiple drug combinations (Project 2).

Conclusions

Core B continued to provide statistical analysis and data management support for all research projects in the IMPACT study.

Core C: Pathology Core
(Director: Ignacio Wistuba, M.D.)

The IMPACT interdisciplinary research proposal for studying targeted therapy of lung cancers has required extensive histopathologic, IHC, and molecular studies of cell and tissues specimens, which have been assisted, coordinated or performed by the Pathology Core. One of the most important roles of the Pathology Core has been to provide professional technical services for proper procurement, storage and use of human and animal tissues, as well as technical assistance for IHC analysis. In addition, the Pathology Core has provided assistance for collection and evaluation of tissue specimens in IMPACT clinical trials in lung cancer patients.

Aim 1   Develop and maintain repository of tissue, cell and serum specimens from patients with lung neoplasia, as requested by the various component projects.

Summary of Research Findings
During the current project period, we continued the collection and processing of tissue and cell specimens for the IMPACT clinical trials.

Project 1. For the IMPACT clinical trial using erlotinib (Tarceva^R), chemotherapy, and radiotherapy in advanced non-small cell lung cancer (NSCLC) patients (Protocol 2005-1023; Principal Investigator: R. Komaki), the Pathology Core assisted in the evaluation of tumor tissue specimens of 29 NSCLC cases for biomarker analysis. Of those, 15 cases have been considered adequate for analysis and the other 6 cases were previously examined for *EGFR* (exon 18-21) and *KRAS* (codons 12, 13 and 61) mutations, *EGFR* copy number by fluorescent *in situ* hybridization (FISH), and EGFR protein (total and phosphorylated [p]) expression by immnuhistochemistry (IHC). No *EGFR* and *KRAS* mutations were detected in the remaining 8 cases, and one case showed increased *EGFR* copy number (high polysomy) in this set. Currently, we are preparing the samples to complete the analysis in the 8 remaining cases. Based on preclinical data provided by investigators of Project 1, we will include the analysis of the expression by IHC of epithelial-mesenchymal transition (EMT) markers, such as E-cadherin and vimentin.

Project 2. For the IMPACT clinical trial using ZD6474 in NSCLC patients with malignant pleural effusion (Protocol 2005-9029; Principal Investigator: C. Jimenez), we have collected, processed, and banked specimens from the effusions from 19 patients at baseline and several time points (Table 1). Multiple samples have been prepared from those specimens, including the following: a) frozen cell pellet; b) formalin-fixed and paraffin-embedded (FFPE) pellets cell blocks and histology sections; c) alcohol-fixed and frozen cell smears; d) cell pellets in RNA later; and, e) cell pellets frozen in DMSO 12% (Table 4).

Table 4. Number of samples collected and processed from pleural fluid from patients enrolled in the IMPACT clinical trial 2005-9029.

Sample Type	Time Points Weeks)									Total
	0	2	3	4	5	6	10	18	End	
Supernatant	40	32	2	2	2	10	6	2	2	98
Cell Pellets										
Frozen	32	6	8	0	0	6	8	0	0	60
RNA Later	20	15	1	1	1	6	3	1	0	48
Cell Block	24	11	1	1	1	5	3	1	1	48
Smear	67	48	4	4	4	16	6	4	4	157

Aim 2   Develop innovative tissue and cell reagents from lung cancer patients for the investigation and validation of the molecular endpoints relevant to each component project.

Summary of Research Findings
The development of new tissue and cell reagents from lung cancer patients was completed in 2008. These reagents and data are available for future analysis of molecular abnormalities of lung cancer, including tissue microarrays (TMA), lung cancer cell lines repository, lung cancer heterotransplants (in collaboration with L. Mao), and pleural fluid specimens from lung cancer patients (in collaboration with C. Jimenez). In particular, the repository of pleural fluid specimens from 120 patients have been fully characterized from the cytology standpoint and the clinical information has been annotated.

Aim 3   Process human and animal cell and tissues for histopathological, immunohistochemical (IHC) and molecular analyses, including tissue microdissection, as required by each component project.

Summary of Research Findings
As indicated in the other Aims, most of the collaborations with the IMPACT research projects have been completed as reported previously. The preparation and processing of tissue specimens for research projects are described in Aim 4 and includes mostly histopathological, IHC, and molecular analysis.

Aim 4   Perform and evaluate IHC analysis in human and animal cell and tissue specimens, as required by the various component projects.

Summary of Research Findings
Most of our collaborations with the other IMPACT projects were completed on 2007 and 2008. During the current project period we have collaborated with the IMPACT investigators to finalize the preparation of manuscripts for publications (see below, *a* to *c*). In addition, we have finalized most of our own research activities (see below, *d* to *f*).

Completed or published projects during the current project period in collaboration with IMPACT research projects.

a) Analysis of GRP78, IL-11R and Eph5A markers in human NSCLC tissue specimens (Collaboration w ith P roject 3, R . P asqualini). Fr om thi s w ork, tw o m anuscripts ar e i n preparation**.**

b) Molecular abnormalities (protein expression and methylation) analysis of *TCF21* gene in NSCLC (Collaboration with Project 6, R. Krahe)**.** From this work, one manuscript is in preparation.

c) Analysis of the IHC expression of angiogenesis-related markers HIF-1α and carbonic anhydrase IX (CA IX) in tumor specimens from 330 NSCLC patients using TMAs (Collaboration with Project 1, R. Herbst)**.** These data hav e been s hared with Dr. J. Heymach's lab (M.D. Anderson Thor acic/Head and N eck M edical O ncology D epartment) an d hav e bee n used to ex pand thei r *in vitro* mechanistic w ork on the r ole of H IF-1α pathway i n l ung c ancer progression and response to therapy. This work was reported last year.

d) Association of *EGFR* gene abnormalities with estrogen and progesterone receptors expression. A paper by G. Raso et al. was published in 2009 in *Clinical Cancer Research*. The submission of this paper was reported last year.

e) Characterization of HER family receptors markers and *EGFR* gene abnormalities in NSCLC primary tumors and brain metastasis. A paper by M. S un et al . w as publ ished in 2009 in *Clinical Cancer Research*. The resubmission of this paper was reported last year.

f) Analysis of VEGFA and VEGFR-2 copy number and mutation in NSCLC. In collaboration with D r. R . Herbst (Project 1) , w e pr eviously s howed that V EGFR and E GFR pat hways ar e positively correlated in early stage NSCLC and IHC expression of p-VEGF-R2 is an indicator of worse overall survival i n stage I-IIIA N SCLC. Recently, w e extended our study to characterize the frequency of VEGFR2 gene (*KDR*) variations and *KDR* and VEGF copy number gains in NSCLC tumor tissue and correlated with patients' clinicopathologic features, including outcome. We extracted DNA from microdissected tissues obtained from 200 surgically resected NSCLCs. KDR single nucleotide polymorphisms (SNP) 889G/A (rs2305948), 1416A/T (rs1870377), and -37A/G (rs2219471) w ere genoty ped by P CR-based s equencing. *KDR* and *VEGF* gene c opy numbers w ere ex amined by r eal-time quanti tative P CR. Tum or s ections were immunohistochemically s tained an d anal yzed for pr otein expression of V EGFR2, V EGF, and CD34 for m icrovascular density (MVD). NSCLC patients with *KDR* 1416 AT/TT variant genotypes had significantly better overall survival (OS; hazard ratio [HR] = 0.56; P = 0.035) than did patients with the wild-type *KDR* 1416 AA genotype. NSCLC patients who were treated with adjuvant c hemotherapy and w ho h ad the *KDR* 1416 A T/TT (P = 0.01 5) or -37 AG /GG (P = 0.029) variant genotypes showed significantly better OS than did the r est of the pat ients in the study population. *KDR* and *VEGF* gene copy gains were detected in 34 (37%) and 2 (2%) of 91 NSCLC tumors, respectively. *KDR* gene copy gain was associated with a worse OS in NSCLC patients (HR= 2.96; P = 0.004). Furthermore, tumors with *KDR* gene copy gain had significantly higher VEGFR2 expression, lower VEGF expression, higher MVD, and larger vessel areas than did tum ors l acking *KDR* gene c opy gai n. O ur fi ndings suggest an as sociation bet ween *KDR* SNP genot ypes and s urvival i n NSCLC pati ents, i ncluding thos e pati ents r eceiving adj uvant chemotherapy. *KDR* copy number gain was frequently identified in NSCLC and was associated with a worse survival in these patients, and with tumors' increased angiogenesis. From this work, an abstract was presented as poster in the 13th World Lung C ancer Conference (IASLC), San Diego, CA (July 31-August 4, 2009).

g) Role of *NKX2-1* (*TITF-1*) gene amplification and protein overexpression in lung cancer and its association with *EGFR* abnormalities. Following our data i ndicating that *NKX2-1* amplification by FIS H correlated with outc ome i n pati ents w ith NSCLC, i ncluding bot h adenocarcinoma and squamous cell carcinoma histology, we have expanded our initial observation to a l arger s et of N SCLC c ases and es tablished a D NA q -PCR methodology for analysis of *NKX2-1* copy number analysis. This work was presented as a platform presentation in the 100 A nnual AACR Meeting, Denver, CO (April 1 8-22, 2009). We have al so investigated the role of *NKX2-1* methylation in the silencing of expression in squamous cell carcinomas of the lung (presented as a poster in the IASLC in 2009). A manuscript with all these data is in the final stages of preparation and i t will be s ubmitted in March 2010 by X. Tang and H . Kadara et al.

h) *EGFR* and *HER2* copy analysis by FISH in tumor specimens from NSCLC patients, and correlation with *EGFR* and *KRAS* mutations and patients' clinicopathologic features. During the l ast year we performed the s tatistical analysis of the gene c opy number by FISH of these two genes in a large series of NSCLC (n=330) tumors. Data from the preliminary analysis were presented as a poster discussion at the IASLC in 2009, and a manuscript is in preparation by Sun et al.

Key Research Accomplishments

- Maintained a large repository of l ung cancer tissue, cytology, and cell lines specimens with annotated clinical data, to be utilized for research projects and clinical trials.
- Maintained a series of lung cancer heterotransplants in mice in collaboration with Project 6.
- Characterized the m olecular abno rmalities of *VEGF* and *KDR* (VEGFR2) genes i n NSCLC, including the identification of high frequency of *KDR* copy number gain (~30%) in tumors.
- Published 2 papers in a prestigious peer-reviewed journal (*Clinical Cancer Research*) on the characterization of E GFR-related abnormalities in NSCLC brain metastasis and on the expression of ER expression in *EGFR*-mutant and wild-type tumors.
- Presented 4 abstracts a t s cientific meetings: one at the 2 009 A ACR annual m eeting, three at the 13th IASLC World Lung Cancer Conference (IASLC), and with an additional abstract to be presented at the 2010 AACR annual meeting.

Conclusions

The P athology C ore ha s as sisted and c ollaborated ac tively w ith s everal r esearch pr ojects to perform multiple histopathological, immunohistochemical, and genetic studies in a large series of lung cancer tissue, including the collection and processing of prospectively collected samples from two ongoing clinical trials. In addi tion, the Pathology Core has managed to c omplete and publish several research activities, which fully integrate with some of the IMPACT research projects. The P athology C ore has s uccessfully fulfilled the goals pr oposed for the fi fth year of the IMPACT program.

Core D: Imaging Core: Provide Imaging Support for IMPACT Projects
(PI and co-PI: Juri Gelovani, M.D., Ph.D.; Chun Li, Ph.D.)

Summary of Research Findings

As reported in the last annual report, Core D completed the work proposed for Projects 2 and 3 related to the dev elopment and e valuation of s tandard oper ating pr ocedures and c hemistry, manufacturing, and control sections of the INDs for the Phase I clinical studies with the novel agents explored in those projects. The final results of these studies are described in the reports for P rojects 2 and 3.   The   contracted G LP s tudies on tox icology of F -PEG6-IPQA were performed by the Charles River Laboratories (MA) as also reported in Project 2.

As C ore D s ervices w ere no l onger r equired by the   Projects, w e hav e w orked to as sist investigators i n pr eparing any r elated publ ications and   grant s ubmissions bas ed on w ork performed in the core over the past year.

Key Research Accomplishments

- Completed microPET/CT imaging (and autoradiography) studies in mice with orthotopic models of different NSCLC using ^{18}F-PEG6-IPQA in Project 2.
- Completed microPET imaging studies for Project 3.
- Completed synthesis and ^{64}Cu/^{111}In radiolabeling of novel cyclic nanopeptides for animal imaging studies in Project 3.

Conclusions

The Imaging Core provided and completed highly specialized imaging support as requested by the project leaders of IMPACT for their designated research projects.

DRP-1: Treatment of Malignant Pleural Effusion with ZD6474, a Novel VEGFR and EGFR TK Inhibitor
(PI and co-PI: Roy Herbst, M.D., Ph.D., Carlos Jimenez, M.D.)

Aim 1   To determine clinical effect of ZD6474.

Aim 2   To investigate biological correlates.

Aim 3   To investigate radiographic correlates.

Aim 4   To assess quality of life.

Summary of Research Findings

The am ended single ar m, open-label study to ev aluate the eff icacy of ZD 6474 on th e management of pleural effusion in NSCLC patients was activated in June of 2007.

To date, 27 patients have been enrolled. Eight patients are excluded from analysis:

- Seven patients did not receive medication:

  - Non-compliance (1)

- Benign pleural effusion (2)

- Non-amenable for intrapleural catheter (IPC) placement (3)

- Renal dysfunction (1)

- One patient was not evaluable due to placement of a defective intrapleural catheter.

Best response

- Twelve patients had stable disease

- One patient had partial response

- Four patients had disease progression

- One patient had treatment failure due to emphysema

- One patient enrolled recently with no information

Days with intrapleural catheter in place
Ten patients have had their intrapleural catheters removed within 28 days after insertion. The last patient still has the intrapleural catheter in place after 126 days.

Related adverse events
One patient stopped the medication at week 6 after developing neurological symptoms and hyponatremia. One patient with a defective intrapleural catheter was excluded from the study at week 6. One patient stopped the medication at week 7 due to recurrence of QTc prolongation after dose reduction. One patient has the IPC dislodged on day 15, and the IPC was later removed on day 23 due to emphysema.

Correlative studies
In collaboration with Dr. Wistuba, Core C Director, we have been collecting pleural effusion specimens including cryopreserved cell pellets, supernatants, FFPE blocks, smears, and RNA for correlative studies (see Table 2, Core C). In addition, serial pleural fluid collection of ten patients with NSCLC and intrapleural catheters not receiving ZD6474, will be analyzed as a control group.

Key Research Accomplishments

- Enrolled 27 patients to the clinical trial.
- Collected specimens for correlative analyses from all patients (4 pleural fluid samples and four blood samples per patient).

Conclusions

We continue to focus efforts on completing accrual of patients to the trial, and all analyses will be performed after the trial is completed.

DRP-2: TALK - Teens and Young Adults Acquiring Lung Cancer Knowledge

(PI: Alexander V. Prokhorov, M. D., Ph.D.)

Ninety percent of lung cancer cases in adults are direct results of smoking. In children and young adults, tobacco use remains a major public health problem in spite of the recent declines in smoking prevalence among children and adolescents. Over the past 2-3 decades, numerous factors of smoking initiation among adolescents have been thor oughly i nvestigated. A considerable volume of literature is currently available providing important clues with respect to designing tobacco prevention and cessation among youth.

Focusing on this major public health problem – tobacco use among young individuals and lack of in-depth knowledge of lung cancer issues – Project TALK (Teens and Young Adults Acquiring Lung Cancer Knowledge) was conceived and funded as a s moking cessation/prevention pilot project for c ulturally diverse hi gh-risk y oung popul ations that i nclude s chool dr op-outs, economically disadvantaged, and underserved. Using modern technologies, the Departments of Behavioral Science and Thoracic/Head & Neck Medical Oncology have joined their efforts to conduct this developmental project under the l eadership of Dr. Alexander V. Prokhorov. The project will assist in making major advances in lung cancer education and prevention among youth. Project TALK will produce a CD-ROM-based education/behavior change for teenagers and young adults (15-24 years of age).

We have thus been devoting our effort in 4 tasks as described in the Statement of Work based on the project timeline:

Aim 1 **Develop intervention program.** Focus groups will be held with adolescents and young adults to ensure we are capturing the essence of the program, using the right messages, and employing the appealing video and animated characters. (Years 1-2)

Summary of Research Findings
This aim was completed and summarized in previous reports.

Aim 2 **Develop and beta-test CD-ROM.** This i ncludes the des ign of the ani mation, illustrations, scripts and accompanying videos. (Years 1-2)

Summary of Research Findings
This aim was completed and summarized in previous reports.

Aim 3 **Implement program in agreed upon locations and recruit young adults to participate in the study.** (Years 3-4)

Summary of Research Findings
This aim was completed and summarized in previous reports.

Aim 4 **Collect and analyze data.** (Years 3-4)

Summary of Research Findings
This aim was completed and summarized in previous reports.

Melissa Kaime, MD, CAPT, MC, USN, Deputy Director of the Peer Reviewed Medical Research Program encouraged Dr. Alexander Prokhorov to apply for a Peer Reviewed Medical Research

Program grant after seeing the "Escape with Your Life" educational video game developed in Project TALK. In April 2009, Dr. Prokhorov was awarded a Peer Reviewed Medical Research Program grant of $3.7 million to develop a tobacco prevention and cessation video game intervention for U.S. service members at Fort Hood, TX, based on the "Escape with Your Life" prototype. Of 40 nutrition and health promotion proposals submitted, Dr. Prokhorov was the only investigator to receive DoD funding for this purpose. This new study is titled "Project COMBAT," and video game development is well underway after conducting a site visit to Fort Hood in May 2009.

In addition, Dr. Prokhorov and colleagues were invited to attend and present at the Military Health Research Forum in Kansas City, MO on September 3, 2009. The presentation, "Behavioral Therapies for Tobacco Control" included an introduction to tobacco use in the military, a description of Project TALK's "Escape With Your Life" video game and an overview of Project COMBAT.

Key Research Accomplishments

- Completed all proposed aims.
- Received a Peer Reviewed Medical Research Program award for $3,700,000 for an expanded Project TALK-based research program.
- Conducted a site visit to Fort Hood, TX to collect data for the new "Escape With Your Life" video game.

Conclusions

This project was completed as reported in the previous annual report. We have used the results from this project to obtain new funding for the "Project COMBAT" research that began in April 2009. We hope to continue to make major advances in lung cancer education and prevention among youth.

Career Developmental Project (CDP1): Identification of Membrane Proteins in Bronchial Epithelia Cells as Biomarkers of Early Detection for Lung Cancer

(PI: Ja Seok Peter Koo, Ph.D.)

Lung cancer is the leading cause of cancer deaths. Early detection of the malignant lesion leads to an improved 5-year survival rate after surgical resection. Therefore, advanced screening tools are needed urgently to detect lung cancer at an early stage to improve control of such deadly lung cancer.

It should be noted that Dr. Kang relocated to another institution after the beginning of the current project period. Dr. Ja Seok Koo has been appointed by the IMPACT Principal Investigator, Dr. Waun Ki Hong, to assume ongoing research of this project.

Aim 1   To isolate membrane proteins uniquely expressed on the surface of squamous metaplasia using organotypically cultured bronchial epithelial cells.

Summary of Research Findings
This aim was completed as reported in the previous annual report.

Aim 2   To identify differentially represented proteins using proteomics.

Summary of Research Findings
The goal of this study is to identify and characterize cell surface markers specific for lung cancer cells, as these proteins have strong potential for novel diagnostic and therapeutic biomarkers.

Our pr evious attem pt to i dentify membrane pr oteins s pecific for cancer cells was not successful. Several te chnical i ssues w ere involved. H ere, w e r epeated 2D - Proteomics M ass s pectrometry analysis of the two protein spots (indicated b y r ed ar rows i n the 2D gel) us ing optimized new conditions (Figure 18). After extensive database s earching, w e found tha t the protein in the lower spot turn out to be an unk nown pr otein and identified th e pr otein i n the upper spot as a receptor for activated kinase C (RACK1).
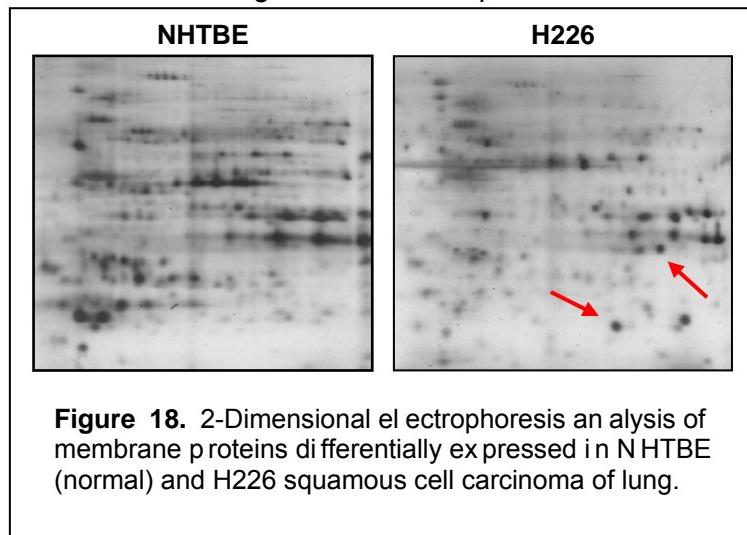
Figure 18. 2-Dimensional el ectrophoresis an alysis of membrane p roteins di fferentially ex pressed i n N HTBE (normal) and H226 squamous cell carcinoma of lung.

RACK1 is a cytosolic membrane protein and also known as lung cancer oncogene 7 and PIG21 (proliferation-inducing gene 21). Its official gene name is GNB2L1 (guanine nucleotide binding protein (G protein), beta polypeptide 2-like 1).

RACK1 was originally identified as an anchoring protein for protein kinase C β (PKC β), which it stabilizes i n the ac tive s tate and anc hors to m embranes or func tional s ites. R ACK1 al so interacts with several other important signaling proteins including the Src kinase family, integrin β1, - β2, -β3, and β5, beta-spectrin and dynamin, RasGAP, the andr ogen r eceptor, i nsulin-like growth fac tor 1 r eceptor (IGF-1r), E pstein-Barr v irus t rans-activator protein B ZLF1, p73 , and pRB. This interaction suggests that RACK1 may act as an anchor or adaptor protein, recruiting other proteins to various transmembrane receptors, providing a platform for protein-protein interactions and ac ting as the foc us for several cell-signaling pathways. Accordingly, a num ber of cellular functions have been attributed to R ACK1, such as cell growth, adhesion, protrusion and c hemotactic m igration. R ACK1 may also contribute t o angi ogenesis and tu mor gr owth, since i t was s hown to b e up-regulated dur ing a ngiogenesis *in vitro* and *in vivo*, and was also expressed in tumor angiogenesis. RACK1 expression was higher in human non-small cell lung and colon carcinomas than in the corresponding normal tissues.

Recent studies [7] also showed that RACK1 is up-regulated in oral squamous carcinoma, and linked to clinical invasiveness and metastasis. In addition, elevated RACK1 expression predicts poor clinical outcome in oral squamous carcinoma, similar to Ki67 [8]. These studies suggest that R ACK1 is a poten tial bi omarker for prognosis and a ther apeutic tar get for squamous carcinoma. We w ill det ermine w hether R ACK1 ex pression i s al tered (overexpressed) i n l ung cancer cells using Western blot analysis. In the meantime, we are collecting samples of various NSCLC cell lines, and the samples will be used for verification of the expression of RACK1. We are aiming to c ollect s amples of at least 30 d ifferent genetically certified lung cancer cell lines. Specificity of R ACK1 expression w ill al so be deter mined by i mmunohistochemical anal ysis using tissue slides.

Aim 3   To verify the differentially represented proteins using PCR, Western blotting, and immunocytochemistry.

Summary of Research Findings

Work on this Aim was dependent upon further verification of the results gained in Aim 2 on the differentially expressed membrane proteins. We will begin these studies to confirm the represented proteins using real-time PCR, Western blotting analysis, and immunohistochemical staining during the next project period following our request for an unfunded cost extension.

Key Research Accomplishments

- Demonstrated that RACK1 (receptor for activated kinase C) may act as an anchor or adaptor pr otein to r ecruit other pr oteins to v arious tr ansmembrane r eceptors and provide a platform for protein-protein interactions.
- Obtained v iable r esults on di fferentially ex pressed m embrane pr oteins, and w ill begin Aim 3 during the next project period.

Conclusions

Two Dimensional-PAGE and Proteomics Mass spectrometry (HPLC-MS-MS) analysis of the two protein s pots that ar e uni quely ex pressed i n s quamous c ell c arcinoma c ell l ines i dentified the spots as a receptor for activated kinase C (RACK1) and one unknown protein. As a result, RACK1 is proposed as a biomarker to detect squamous cell carcinoma. Further characterization is underway and will be illustrated in the next progress report.

KEY RESEARCH ACCOMPLISHMENTS

Project 1: Targeting epidermal growth factor receptor signaling to enhance response of lung cancer to therapeutic radiation.

- Enrolled 38 pati ents onto the erlotinib (Tarceva) plus radiotherapy for locally advanced NSCLC trial, and completed evaluation of 30 of these patients for response and toxicity.
- Discovered the r elationship bet ween the epithelial-to-mesenchymal tr ansition and radiosensitivity of NSCLC cells (preclinical data).
- Demonstrated that pr etreatment with gefi tinib e xerts a r adioprotection of H 1299-CDH1 cells.
- Demonstrated that s mall m olecule i nhibitors of both c -Met and IG F-1R pr oduce a significant radiosensitizing effect on NSCLC cells.
- Completed an as sessment of the c ombination of er lotinib (Tarceva) and r adiation i n a NSCLC xenograft tumor model.

Project 2: Molecular imaging of EGFR expression and activity in targeted therapy of lung cancer

- Contracted for the IND-driven G LP tox icology study of ^{18}F-PEG6-IPQA i n r ats, w hich was completed in this last funded period, enabling us to move the agent into clinical trial.
- Completed Phase I clinical protocol development. This protocol has been reviewed and approved by the D epartment of N uclear M edicine w here tr ial will be c onducted, and i s now pending IRB approval.
- Developed an automated manufacturing process that complies with cGMP based on the lab m anual pr ocess. Wi th thi s cGMP pr oduction, w e are abl e to pr oduce F DA-acceptable imaging agent to be used in the clinical phase I trial.
- Wrote the Inv estigational Drug Application (IND) to be s ubmitted to FD A for approval to conduct the phase I trial.
- Two manuscripts (on studies in mice and non-human primates) are in the final stages of preparation for publication.

Project 3: Targeted peptide-based systemic delivery of therapeutic and imaging agents to lung tumors

- Used labeled targeted peptide motifs themselves as imaging tools.
- Demonstrated that AAVP-based m olecular-genetic imaging appears to be superior to FDG in side-by-side comparisons for predicting therapeutic response.
- Designed and dev eloped nanote chnology-based (i.e., bottom -up, self-assembled) biocompatible networks of phage-gold as nanotechnology-based molecular sensors and reporters.

Project 4: Inhibition of bFGF Signaling for Lung Cancer Therapy

- Discovered that blocking the FGFR1 signaling by DNFR1 leads to inhibition of cell growth by blocking cells at the G2 phase of the cell cycle through decreased expression, phosphorylation, and k inase func tion of p34c dc2 c ombined w ith i ncreased Wee1, and the cyclin-dependent kinase inhibitors p21 and p27.

- Determined that Geldanamycin inhibits the growth of lung cancer cells by a G2/M block and by a mechanism that in part emulates the blocking of FGFR signaling by DNFR1 expression.
- Demonstrated the ability of bFGF to increase Hsp90 level and the ability of the Hsp90 inhibitor Geldanamycin to inhibit FGF signaling, suggesting that Hsp90 is a positive mediator of FGF effects.

Project 5: Targeting mTOR and Ras signaling pathways for lung cancer therapy

- Demonstrated that concurrent treatment with rapamycin and LY294002 is more effective than sequential treatments with either rapamcyin followed with LY294002 or LY294002 followed by rapamcyin in inhibiting the growth of cancer cells.
- Discovered that the combination of sub-optimal concentrations of RAD001 and BEZ235 exhibited synergistic effects on inhibiting the growth of the lung cancer cells evaluated.
- Enrolled a total of 11 patients to the study, including one patient on the control arm.

Core B: Biostatistics & Data Management Core

- Continued to provide statistical support in the clinical trials for Project 1 and DRP-1.
- Provided statistical support for Projects 2, 3, 4, 6, and Pathology Core.
- Continued to work closely with the Project 4 (PI: Dr. Reuben Lotan) on synergy studies of combination drug treatment in cell lines to determine whether the effect is synergistic, additive, or antagonistic.
- Developed methods and provided codes to construct the confidence interval for the interaction index for the Emax model with Dr. Lotan in Project 4. Further enhancements of estimating the parameters of the Emax model were provided by interpolation and extrapolation to improve the model convergence for cases when the observed data points do not yield adequate model fit. Sensitivity analysis was implemented for checking the model robustness as well.
- Developed a Bayesian dose-finding trial design with multiple drug combinations (Project 2).

Core C: Pathology Core

- Maintained a large repository of lung cancer tissue, cytology, and cell lines specimens with annotated clinical data, to be utilized for research projects and clinical trials.
- Maintained a series of lung cancer heterotransplants in mice in collaboration with Project 6.
- Characterized the molecular abnormalities of *VEGF* and *KDR* (VEGFR2) genes in NSCLC, including the identification of high frequency of *KDR* copy number gain (~30%) in tumors.
- Published 2 papers in a prestigious peer-reviewed journal (*Clinical Cancer Research*) on the characterization of EGFR-related abnormalities in NSCLC brain metastasis and on the expression of ER expression in *EGFR*-mutant and wild-type tumors.
- Presented 4 abstracts at scientific meetings: one at the 2009 AACR annual meeting, three at the 13th IASLC World Lung Cancer Conference (IASLC), and with an additional abstract to be presented at the 2010 AACR annual meeting.

Core D: Imaging Core

- Completed microPET/CT imaging (and autoradiography) studies in mice with orthotopic models of different NSCLC using ^{18}F-PEG6-IPQA in Project 2.
- Completed microPET imaging studies for Project 3.
- Completed synthesis and ^{64}Cu/^{111}In radiolabeling of novel cyclic nanopeptides for animal imaging studies in Project 3.

DRP-1: Treatment of Malignant Pleural Effusion with ZD6474, a Novel VEGFR and EGFR TK Inhibitor

- Enrolled 27 patients to the clinical trial.
- Collected s pecimens for c orrelative anal yses fr om al l pati ents (4 pl eural fl uid s amples and four blood samples per patient).

DRP-2: TALK - Teens and Young Adults Acquiring Lung Cancer Knowledge

- Completed all proposed aims.
- Received a P eer R eviewed M edical R esearch P rogram a ward for $3,700,000 for the Project TALK research program.
- Conducted a site visit to Fort Hood, TX to collect data for the new "Escape With Your Life" video game.

CDP1: Identification of Membrane Proteins in Bronchial Epithelia Cells as Biomarkers of Early Detection for Lung Cancer

- Demonstrated that RACK1 (receptor for activated kinase C) may act as an anchor or adaptor pr otein to r ecruit other pr oteins to v arious tr ansmembrane r eceptors and provide a platform for protein-protein interactions.
- Obtained v iable r esults on di fferentially ex pressed m embrane pr oteins, and w ill begin Aim 3 during the next project period.

REPORTABLE OUTCOMES

Publications *(attached in Appendix)***:**

__2008 Publications Previously Reported without PMID Numbers:__

Behrens C, Lin HY, Lee JJ, Raso MG, Hong WK, Wistuba II, Lotan R. Immunohistochemical expression of basic fibroblast growth factor and fibroblast growth factor receptors 1 and 2 in the pathogenesis of lung cancer. *Clinical Cancer Research.* 2008 Oct 1;14(19):6014-22. PMID: 18829480.

Cardó-Vila M, Zurita AJ, Giordano RJ, Sun J, Rangel R, Guzman-Rojas L, Anobom CD, Valente AP, Almeida FC, Lahdenranta J, Kolonin MG, Arap W, Pasqualini R. A ligand peptide motif selected from a cancer patient is a receptor-interacting site within human interleukin-11. *PLoS ONE.* 2008;3(10):e3452. PMCID: PMC2565473.

Colella S, Richards KL, Bachinski LL, Baggerly KA, Tsavachidis S, Lang JC, Schuller DE, Krahe R. Molecular signatures of metastasis in head and neck cancer. *Head & Neck.* 2008 Oct;30(10):1273-83. PMID: 18642293.

Ozawa MG, Zurita AJ, Dias-Neto E, Nunes DN, Sidman RL, Gelovani JG, Arap W, Pasqualini R. Beyond receptor expression levels: the relevance of target accessibility in ligand-directed pharmacodelivery systems. *Trends in Cardiovascular Medicine.* 2008 May;18(4):126-32. PMID: 18555185.

Staquicini FI, Tandle A, Libutti SK, Sun J, Zigler M, Bar-Eli M, Aliperti F, Pérez EC, Gershenwald JE, Mariano M, Pasqualini R, Arap W, Lopes JD. A subset of host B lymphocytes controls melanoma metastasis through a melanoma cell adhesion molecule/MUC18-dependent interaction: evidence from mice and humans. *Cancer Research.* 2008 Oct 15;68(20):8419-28. PMCID: PMC2718575.

Tang X, Varella-Garcia M, Xavier AC, Massarelli E, Ozburn N, Moran C, Wistuba II. EGFR Abnormalities in the Pathogenesis and Progression of Lung Adenocarcinomas. *Cancer Prevention Research.* 2008 Aug;1(3):192-200. PMID: 19138956.

Wang X, Hawk N, Yue P, Kauh J, Ramalingam SS, Fu H, Khuri FR, Sun S-Y. Overcoming mTOR inhibition-induced paradoxical activation of survival signaling pathways enhances mTOR inhibitors' anticancer efficacy. *Cancer Biology & Therapy.* 2008 Dec;7(12):1952-8. PMCID: PMC2762753.

Wang X, Yue P, Kim YA, Fu H, Khuri FR, Sun SY. Enhancing mammalian target of rapamycin (mTOR)-targeted cancer therapy by preventing mTOR/raptor inhibition-initiated, mTOR/rictor-independent Akt activation. *Cancer Research.* 2008 Sep 15;68(18):7409-18. PMCID: PMC2562339.

Zhang S, Schafer-Hales K, Khuri FR, Zhou W, Vertino PM, Marcus AI. The tumor suppressor LKB1 regulates lung cancer cell polarity by mediating cdc42 recruitment and activity. *Cancer Research.* 2008 Feb 1;68(3):740-8. PMID: 18245474.

2009 – 2010 Publications:

Giordano R J, Cardó-Vila M, Salameh A, Anobom C D, Zeitlin B D, Hawke DH, Valente A P, Almeida FC, Nör JE, Sidman RL, Pasqualini R, Arap W. From combinatorial peptide selection to dr ug pr ototype (I): Targeting the v ascular endothel ial gr owth fac tor r eceptor pathw ay. *Proceedings National Academy of Science U S A.* 2010 Feb 26. PMID: 20190181

Kong M , L ee J J. A pplying E max model and Bivariate T hin P late S plines to A ssess D rug Interactions. Frontiers in Biosciences, (Elite Ed). 2010 Jan 1;2:279-92.PMID: 20036878.

Lee JJ, Lin HY, Liu DD, Kong M. Applying Emax model and interaction index for assessing drug interaction in combination s tudies. Fr ontiers i n B iosciences, (Elite E d) 2010 Jan 1;2:582 - 601.PMID: 20036904.

Lewis V O, Ozawa MG , Deavers M T, Wang G , Shintani T , Arap W , Pasqualini R . The interleukin-11 receptor alpha as a candidate ligand-directed target in osteosarcoma: consistent data from cell lines, orthotopic models, and human tumor samples. *Cancer Research.* 2009 Mar 1;69(5):1995-9. Epub 2009 Feb 24. PMID: 19244100.

Raso MG, Behrens C, Liu S, Prudkin L, Ozburn NO, Woods DM, Tang T, Mehran RJ, Moran C, Lee JJ, and Wistuba I. Immunohistochemical expression of estrogen and progesterone receptors i dentifies a s ubset of non -small c ell l ung c arcinomas and c orrelates w ith *EGFR* mutation. *Clinical Cancer Research* 2009 Sep 1;15(17):5359-68. PMID: 19706809.

Richards KL, Zhang B, Baggerly KA, Colella S, Lang J C, Schuller DE, and K rahe R. Genome-wide hypomethylation i n head and neck c ancer i s m ore pr onounced i n H PV-negative tum ors and is associated with genomic instability. *PLoS One* 2009;4(3):e4941. PMCID: PMC2654169.

Sun S, Behrens C, Feng L, O zburn N, Tang X, Yin G, Komaki R, Varella-Garcia M, Hong WK, Aldape KA, and Wistuba I. HER family receptor abnormalities in lung cancer brain metastases and c orresponding pr imary tu mors. *Clinical Cancer Research* 2009 A ug 1;15(15):4829-37. PMID: 19622585.

Wang X, Sun SY. Enhancing mTOR-targeted cancer therapy. *Expert Opinion on Therapeutic Targets* 2009 Oct;13(10):1193-203. PMCID: PMC2810099.

Manuscripts submitted, in revision, or in review

Richards KL, Zhang B, Sun M, Dong W, Churchill J, Biachinski LL, Yin G, Hayes DN, Wistuba II, Krahe R. Methylation of the candidate biomarker *TCF21* is very frequent across a spectrum of early stage non-small cell lung cancers. *In Revision.*

Abstracts *(attached in Appendix)***:**

Tang X, Liu D, Behrens C, He D, Sun M, Rice D, Lee JJ, H ong WK, Wistuba II. *TITF-1* and *EGFR* gene copy variations are associated with prognosis for the patients with non-small cell lung cancer. Proceedings of the 100 th Annual Meeting of th e American Association for Cancer Research; 2009 Denver, CO. Abstract 3841.

Tang X, Yang F,Huang J,Woods D, Corvalan A, Wistuba II. TITF-1 Protein Expression Associates with Gene Methylation and Gene Copy Gain in Non-Small Cell Lung Carcinoma. 13th World Lung Cancer Conference, IASLC, San Francisco, California, July-August, 2009.

Yang F, Tang X, Corvalan A, Behrens C, Lin H, Raso MG, Lee JJ, Minna JD, Wistuba II. Analysis of C opy Number G ain of *VEGF and VEGFR2/KDR* in N on-Small C ell Carcinoma (NSCLC) Cancer Cell Lines and Primary Tumors. 13th World Lung Cancer Conference, IASLC, San Francisco, California, July-August, 2009.

CONCLUSIONS

Project 1: We conclude that the epithelial-to-mesenchymal transition (EMT) plays a significant role in governing not just the intrinsic radiosensitivity of NSCLC cells, but also their sensitivity to inhibitors of the epidermal growth factor receptor (EGFR) and the ability of such inhibitors to radiosensitize these cells. It would be useful to assess the EMT status of patients treated with these combinations. In spite of this finding, results suggest that such combinations might be useful in the clinic. In addition, we conclude that targeting other growth factor receptors such as the c-Met and IG-F1R receptors may be an alternative strategy to using EGFR inhibitors.

Project 2: Development of ^{18}F-PEG6-IPQA for the proposed clinical trial has been largely completed.   We will next pursue IRB and FDA approval of the protocol and IND, respectively, before initiating the clinical phase I trial, which is anticipated during the next unfunded period (no-cost extension).

Project 3: The central working hypothesis in our program is that differential protein expression in the human vascular endothelium associated with lung cancer offers the potential for developing novel diagnostic, imaging, and therapeutic strategies. In essence, combinatorial library selections (peptide- and antibody-based) are leveraged to discover, validate, and exploit the vascular biochemical diversity of endothelial cell surfaces towards a new vascular-targeted pharmacology. Such targeting technologies may lead to the development of ligand-directed agents for application in the treatment of cancer patients. Translational applications, such as first-in-human clinical trials, have now began within the institution, as the Food and Drug Administration (FDA) has recently granted a "safe-to-proceed" status for the first vascular-targeted Investigational New Drug, discovered, developed and being evaluated in patients at MDACC. Such trials will ultimately determine the value of this strategy. Two other drugs are in pre-IND stage and several others in pre-clinical laboratory phase. Long-term, the broader vision of the research is a large-scale mapping of receptors in human vasculature towards a new ligand-directed pharmacology.

Project 4: Our results indicate that the adenoviral vector containing a dominant negative FGFR1 receptor construct, which inhibits FGFR signaling, could potentially be useful for treatment of lung cancer *in vivo.* Studies targeting lung cancer with this adenoviral construct are warranted. Such studies could focus on adenovirus delivery by inhalation using an aerosolized preparation. In terms of understanding FGFR signaling, our studies highlighted an important role for Hsp90 as a downstream mediator of FGF effects. This suggests that combined targeting of FGFR signaling (e.g., by AdVDNFR1) and Hsp90 function (e.g., by Geldanamycin or other Hsp90 inhibitors) could provide additive or synergistic efficacy.

Project 5: During the previous project period, we continued to evaluate whether the effect of mTOR inhibitors on the growth of human NSCLC cells is enhanced in the presence of a PI3K inhibitor or a MAPK inhibitor.   We demonstrated that the combination of sub-optimal concentrations of RAD001 and BEZ235 exhibited synergistic effects on inhibiting the growth of the lung cancer cells and that the combination of RAD001 and BEZ235 was more effective than sequential combination treatment in inhibiting the growth of cancer cells. The findings confirm the feasibility of conducting 'window of opportunity' studies as a novel mechanism to understand the molecular effects on targeted agents on the tumor tissue. We continued to enroll patients to the RAD001 trial to investigate the effect of this drug in operable NSCLC patients and to identify molecular determinants of RAD001 sensitivity and prognosis.

Project 6: This project was completed as reported in the previous annual report.

Biostatistics Core: Core B c ontinued to pr ovide s tatistical analysis and data management support for all research projects in the IMPACT study.

Pathology Core: The Pathology Core has assisted and collaborated actively with several research projects to perform multiple histopathological, immunohistochemical, and genetic studies in a large s eries of l ung cancer ti ssue, i ncluding the c ollection and processing of prospectively collected samples from two ongoing clinical trials. In addition, the P athology Core has managed to complete and publish research activities, which fully integrate with some of the IMPACT research projects. The Pathology Core has successfully fulfilled the goals proposed for the fifth year of the IMPACT program.

Imaging Core: The Im aging C ore pr ovided and c ompleted highly s pecialized imaging support as requested by the project leaders of IMPACT for their designated research projects.

DRP-1: We c ontinue to foc us effor ts on c ompleting ac crual of pati ents to the tr ial, and al l analyses will be performed after the trial is completed.

DRP-2: This project was completed as reported in the pr evious annual report. We have used the results from this project to obtain new funding for the "Project COMBAT" research that began in April 2009. We hope to continue to make major advances in lung cancer education and prevention among youth.

CDP1: Two Dimensional-PAGE and Proteomics Mass spectrometry (HPLC-MS-MS) analysis of the tw o pr otein s pots t hat ar e uni quely ex pressed in s quamous c ell c arcinoma c ell l ines identified the spots as a receptor for activated kinase C (RACK1) and one unknown protein. As a r esult, RACK1 is pr oposed as a bi omarker to detec t squamous c ell c arcinoma. Further characterization is underway and will be illustrated in the next progress report.

REFERENCES

1. Paez, J .G., et al ., *EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy.* Science, 2004. 304(5676): p. 1497-500.
2. Pirker, R. and M. Filipits, *Targeted therapies in lung cancer.* Curr Pharm Des, 2009. **15**(2): p. 188-206.
3. Burris, H .A., 3r d, *Shortcomings of current therapies for non-small-cell lung cancer: unmet medical needs.* Oncogene, 2009. 28 Suppl 1: p. S4-13.
4. Yun, C.H., et al., *Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity.* Cancer Cell, 2007. **11**(3): p. 217-27.
5. Pao, W., et al ., *Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain.* PLoS Med, 2005. **2**(3): p. e73.
6. Perez-Soler, R., *Individualized therapy in non-small-cell lung cancer: future versus current clinical practice.* Oncogene, 2009. 28 Suppl 1: p. S38-45.
7. Wang Z, Jiang L, Huang C, Li Z, Chen L, Gou L, Chen P, Tong A, Tang M, Gao F, Shen J, Zhang Y, Bai J, Zhou M, Miao D, Chen Q. Comparative proteomics approach to screening of potential di agnostic and ther apeutic tar gets for or al s quamous cell c arcinoma. Mol C ell Proteomics. 2008 Sep;7(9):1639-50.
8. Wang Z, Zhang B, Jiang L, Zeng X, Chen Y, Feng X, Guo Y, Chen Q. RACK1, an excellent predictor for poor clinical outcome in oral squamous carcinoma, similar to Ki67.
    Eur J Cancer. 2009 Feb;45(3):490-6.

Appendix

Immunohistochemical Expression of Basic Fibroblast Growth Factor and Fibroblast Growth Factor Receptors 1 and 2 in the Pathogenesis of Lung Cancer

Carmen Behrens,[1] Heather Y. Lin,[2] J. Jack Lee,[2] Maria Gabriela Raso,[3] Waun Ki Hong,[1] Ignacio I. Wistuba,[1,3] and Reuben Lotan[1]

Abstract

Purpose: To identify the patterns of protein expression of basic fibroblast growth factor (bFGF) and FGF receptors 1 and 2 in non-small cell lung carcinoma (NSCLC) and their role in the early pathogenesis of squamous cell carcinoma (SCC) of the lung.

Experimental Design: Archived tissue from NSCLC (adenocarcinoma and SCC; $n = 321$) and adjacent bronchial epithelial specimens ($n = 426$) were analyzed for the immunohistochemical expression of bFGF, FGFR1, and FGFR2, and the findings were correlated with clinicopathologic features of the patients.

Results: High expression of bFGF, FGFR1, and FGFR2 was shown in most NSCLC tumors. The pattern of expression for all markers varied according to tumor histologic type and cellular localization. Cytoplasmic expression scores were significantly higher in tumors than in normal epithelia. Nuclear bFGF ($P = 0.03$) and FGFR1 ($P = 0.02$) levels were significantly higher in women than in men. Although cytoplasmic FGFR1 expression was significantly higher ($P = 0.002$) in ever smokers than in never smokers, nuclear FGFR1 ($P = 0.0001$) and FGFR2 ($P = 0.003$) expression was significantly higher in never smokers. Different prognostic patterns for the expression of these markers were detected for both NSCLC histologic types. Dysplastic changes showed significantly higher expression of all markers compared with squamous metaplasia.

Conclusions: bFGF, FGFR1, and FGFR2 are frequently overexpressed in SCC and adenocarcinoma of the lung. bFGF signaling pathway activation may be an early phenomenon in the pathogenesis of SCC and thus an attractive novel target for lung cancer chemopreventive and therapeutic strategies.

Lung cancer, the leading cause of cancer-related deaths in the United States (1), consists of several histologic types (2), the most common being two non-small cell lung carcinomas (NSCLC): adenocarcinoma and squamous cell carcinoma (SCC; ref. 3). In spite of recent advances, the underlying processes involved in the early pathogenesis of lung cancer remain unclear. NSCLCs are believed to arise after the progression of sequential preneoplastic lesions, including bronchial squamous dysplasias (4). The activation of angiogenesis pathways has been shown to be involved in the development and progression of lung cancer (5–7). A better understanding of the signaling pathways that lead to tumor growth and angiogenesis may help in the development of new and more effective strategies for early detection, targeted chemoprevention, and treatment of lung cancer.

Fibroblast growth factor (FGF) 2, or basic FGF (bFGF), and its transmembrane tyrosine kinase receptors (the FGFRs) make up a large, complex family of signaling molecules involved in several physiologic processes, and the dysregulation of these molecules has been associated with cancer development (8, 9). bFGF belongs to a family of ubiquitously expressed ligands that bind to the extracellular domain of FGFRs, initiating a signal transduction cascade that promotes cell proliferation, motility, and angiogenesis (8–10).

As with some other angiogenesis pathways, the bFGF pathway has been shown to be activated in lung cancer (11–18). Elevated levels of bFGF, FGFR1, and FGFR2 proteins have been detected in NSCLC cell lines (11, 19). Although a few reports discuss the expression of bFGF and FGFRs in NSCLC tumors (12–18), the precise role of these molecules in the early pathogenesis and progression of this tumor is still unknown.

To identify the patterns of expression of bFGF, FGFR1, and FGFR2 proteins in the two major histologies of NSCLC, namely SCC and adenocarcinoma, and their role in the early pathogenesis of SCC of the lung, we investigated the immunohistochemical expression of these three molecules in a large series of archived NSCLC tissue specimens and adjacent

Translational Relevance

A better understanding of the signaling pathways that lead to tumor growth and angiogenesis may help in the development of new and more effective strategies for targeted chemoprevention and treatment of lung cancer. Our findings of frequent overexpression of bFGF, FGFR1, and FGFR2 in NSCLC, and in the early pathogenesis of squamous cell carcinomas, suggest that the activation of the bFGF pathway, which has been proposed to facilitate the development of resistance to antiangiogenic therapy targeting the vascular endothelial growth factor pathway, is an attractive novel target for lung cancer therapeutic and chemopreventive strategies.

lung bronchial epithelial foci and constructed tissue microarray (TMA) specimens. We then correlated our findings with clinicopathologic features of patients with lung cancer.

Materials and Methods

Case selection and TMA construction. For this study, which was approved by our institutional review board, we obtained archived formalin-fixed, paraffin-embedded material from surgically resected lung cancer specimens containing tumor and adjacent lung tissues from the Lung Cancer Specialized Program of Research Excellence tissue bank at The University of Texas M. D. Anderson Cancer Center. Tissue specimens collected between 1997 and 2003 from 321 lung cancer tumors (196 adenocarcinomas and 125 SCCs) were classified by using the 2004 WHO classification system (3) and selected specimens were used for construction of TMAs. In addition, detailed histopathologic semiquantitative analysis of lung adenocarcinoma subtypes (acinar, papillary, solid with mucin, and bronchioloalveolar) using the WHO classification (3) was done in 192 lung adenocarcinomas. After histologic examination, the NSCLC TMAs were constructed by obtaining three 1-mm-diameter cores from each tumor.

To assess the immunohistochemical expression of bFGF, FGFR1, and FGFR2 in the early pathogenesis of NSCLC, we studied formalin-fixed, paraffin-embedded material from 426 specimens of bronchial epithe-

lium surgically resected from 130 patients with NSCLC (mean, 3.3 specimens per patient; range, 1-25 specimens). We histologically classified epithelial lesions by using the 2004 WHO classification system for preneoplastic lung lesions (3). The histologic findings from the epithelia were as follows: normal epithelium (n = 150), basal cell hyperplasia (n = 164), squamous metaplasia (n = 26), and squamous dysplasia (n = 86). The squamous dysplasias were arranged into two groups: low-grade (mild and moderate dysplasias; n = 22) and high-grade (severe dysplasia and carcinoma *in situ*; n = 64). For the epithelial foci TMAs, we used single 2-mm cores in an attempt to capture most of the preneoplastic lesions.

Immunohistochemical staining and evaluation. The following primary antibodies were used for immunohistochemical staining: mouse monoclonal anti-bFGF (BD Biosciences PharMingen; dilution 1:200), rabbit polyclonal anti-FGFR1 (Flg; Santa Cruz Biotechnology; dilution 1:100), and rabbit polyclonal anti-FGFR2 (Bek; Santa Cruz Biotechnology; dilution 1:100). Formalin-fixed, paraffin-embedded tissue sections (5-μm-thick) were deparaffinized and hydrated. For bFGF, antigen retrieval was done by incubating specimens for 5 min in 1% SDS in TBS [100 mmol/L Tris (pH 7.4), 138 mmol/L NaCl, 27 mmol/L KCl] at room temperature. For FGFR1 and FGFR2, antigen retrieval was done by heating specimens in a steamer for 10 min with 10 mmol/L sodium citrate (pH 6.0). We performed protein blocking by incubating specimens for 30 min in 10% bovine serum albumin in TBS with 0.5% Tween 20. Primary antibody incubation was done overnight at 4°C for bFGF and for 2 h at room temperature for FGFR1 and FGFR2.

Specimens were next washed with PBS. They were incubated for 30 min with secondary antibody with Envision Plus Dual Link-labeled polymer (DAKO) and then for 5 min with diaminobenzidine chromogen. Formalin-fixed, paraffin-embedded lung tissues having both tumor and normal tissues were used as a positive control. For a negative control, we used the same specimens that we used for the positive control but replaced the primary antibody with PBS.

The expression was quantified by two observers (C.B. and I.I.W.). Cytoplasmic expression was quantified using a four-value intensity score (0, 1+, 2+, and 3+) and the percentage (0-100%) of the extent of reactivity. Next, the cytoplasmic expression score was obtained by multiplying the intensity and reactivity extension values (range, 0-300). Nuclear expression was quantified using a range of 0 to 100 according to the percentage of positive nuclei present among all tumor or epithelium cells present in the TMA cores.

Statistical analysis. The data were summarized using standard descriptive statistics and frequency tabulations. Associations between the marker expression and patients' clinical demographic variables including age, sex, smoking history, histology type, and pathologic

Table 1. Summary of the clinicopathologic features of lung cancer cases studied

Feature	NSCLC histologic type*		
	SCC (n = 125)	Adenocarcinoma (n = 196)	Total (N = 321)
Mean age (range), y	68.5 (42-90)	64.7 (34-87)	66.3 (34-90)
Sex			
Male	79	74	153
Female	46	122	168
Smoking status [†]			
Never	6	50	56
Ever	118	146	264
TNM stage			
I	68	130	198
II	39	25	64
III	17	34	51
IV	1	7	8

*Values are number of cases unless otherwise indicated.
[†] Smoking status was not available in one patient with SCC.
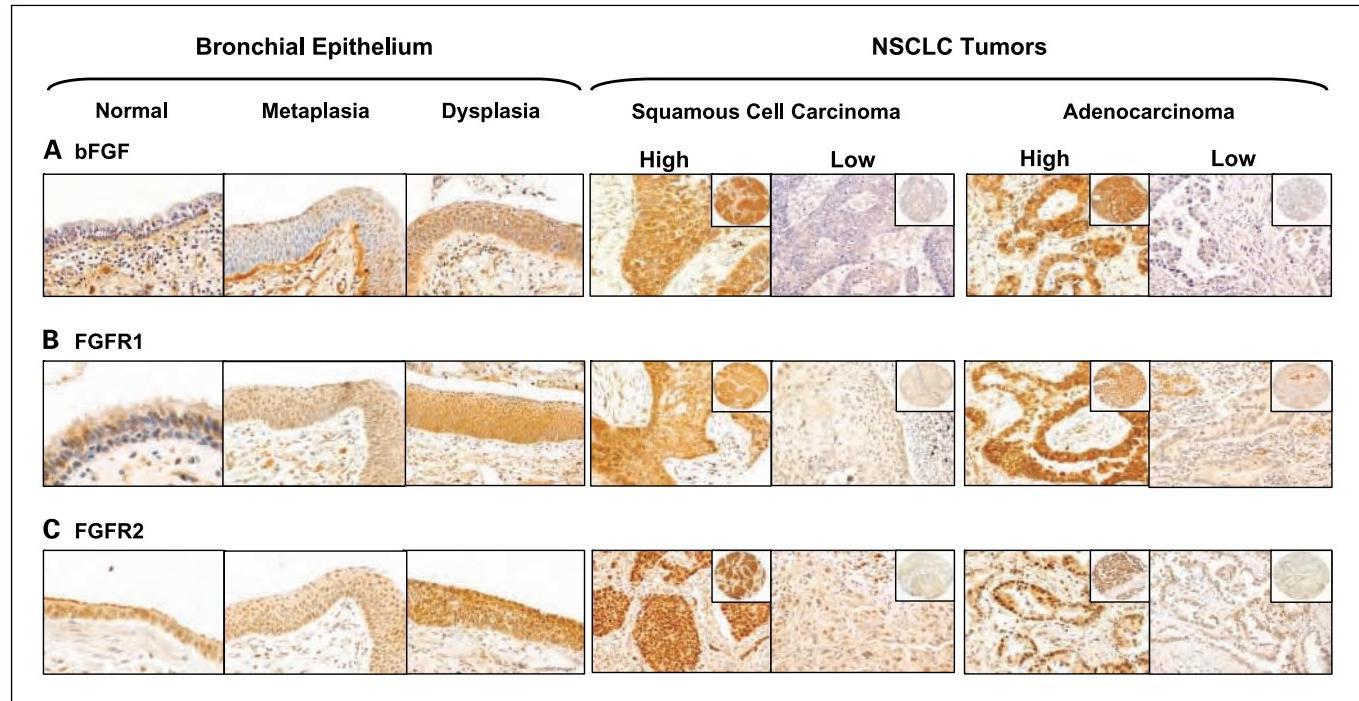
Fig. 1. Microphotographs of immunohistochemical expression of bFGF (*A*), FGFR1 (*B*), and FGFR2 (*C*) in tissue specimens of NSCLC tumor (SCC and adenocarcinoma) and bronchial epithelium (squamous metaplasias, squamous dysplasia, and normal epithelium). For all three markers, immunostaining was preferentially cytoplasmic, but nuclear staining was also detected. In bronchial epithelial specimens, levels of expression were higher in squamous dysplasia than in squamous metaplasia and normal epithelium. For each histologic type and marker in NSCLC tumors, both high and low levels of magnification are shown. *Insets,* tumor TMA core.

stage were assessed using appropriate methods including χ^2 test or Fisher's exact test for categorical variables and Wilcoxon's rank-sum test or Kruskal-Wallis test for continuous variables. Linear mixed-effects model was applied to compare the marker expression level among various histologic progression stage including normal, hyperplasia, metaplasia, dysplasia, and cancer.

We performed overall survival (OS) and recurrence-free survival (RFS) analyses for the expression of bFGF, FGFR1, and FGFR2 by using specimens from 321 NSCLC patients with a median follow-up of 3.78 years. Survival curves were estimated using the Kaplan-Meier method. OS was defined as the time from surgery to death or to the end of the study and RFS as the time from surgery to recurrence or to the end of the study. The effect of the marker expression on patients' OS and RFS was tested for both histologic types separately and adjusted for age, sex, smoking status, and pathologic TNM stage. The binary cutoff points of biomarkers were identified using the Classification and Regression Tree algorithm, in which a cutoff point is determined for each predictor variable such that the two resulting subgroups are the most different in OS. The same cutoffs were used in the RFS analysis. Univariate and multivariate Cox proportional hazards models were used to assess the effect of covariates on OS and RFS. Two-sided *P* values < 0.05 were considered statistically significant. Any significant findings can serve as hypothesis generation and require further confirmation in future studies. All analyses were conducted in SAS 9.1 (SAS Institute) and S-PLUS (Insightful Corp.).

Results

Patients' characteristics. We studied 321 surgically resected lung cancers representing the two major NSCLC histologies using archival tissue specimens. Detailed clinical and pathologic information was available in most cases (Table 1) and included patients' demographic data, smoking history [never smokers or ever smokers (patients who had smoked at least 100 cigarettes in their lifetime)], pathologic TNM staging (20), OS time, and RFS time.

Immunohistochemical expression of bFGF and receptors in NSCLC compared with that in normal epithelium. bFGF, FGFR1, and FGFR2 were detected in the cytoplasm and nuclei of bronchial epithelium and tumor cells (Fig. 1). Overall, NSCLC tumor cells showed higher levels of bFGF, FGFR1, and FGFR2 protein expression than histologically normal bronchial epithelium did. No significant difference in the expression of these three markers was detected in normal epithelium obtained from patients with adenocarcinoma or SCC. Specimens from both histologic types of NSCLC, however, had significantly higher cytoplasmic expression scores for all three markers than did the histologically normal specimens of bronchial epithelium obtained from patients with lung cancer (Figs. 1 and 2). Significant higher FGFR1 and FGFR2 (but not bFGF) nuclear expression scores were also found in tumor than in normal epithelium for patients with adenocarcinoma. On the other hand, significantly higher bFGF (but not FGFRs) nuclear expression score was found in tumor than in normal epithelium in patients with squamous carcinoma (Fig. 2).

Markers of immunohistochemical expression in the sequential pathogenesis of SCC. We investigated the expression of bFGF, FGFR1, and FGFR2 in 426 epithelial specimens containing histologically normal, hyperplastic, squamous metaplastic, or squamous dysplastic bronchial epithelia adjacent to NSCLC obtained from 130 patients. Similar levels of cytoplasmic and nuclear expression were detected in normal and hyperplastic epithelia for all three markers. For cytoplasmic FGFR1 and FGFR2 localizations, squamous metaplastic epithelia showed

significantly lower levels of the markers than did normal (FGFR1, $P = 0.022$; FGFR2, $P = 0.022$) and hyperplastic (FGFR1, $P = 0.047$; FGFR2, $P = 0.049$) epithelia. For nuclear localization, no significant differences in the scores of any markers were detected between normal, basal cell hyperplastic, and squamous metaplastic epithelia. Of interest, squamous dysplastic lesions showed significantly higher levels of expression than did squamous metaplastic lesions for all three markers (Fig. 3). The monotone increase pattern in the expression of these markers from squamous dysplasia to invasive SCC is consistent with the progression model of this tumor type.

Correlation between markers of immunohistochemical expression in NSCLC and clinicopathologic features and disease outcomes. We correlated bFGF and FGFR scores and levels of expression with tumor histologic characteristics, age, sex, smoking history, and TNM pathologic stage. The expression of these markers at the cytoplasmic level was similar in both NSCLC histologic types, except for FGFR2, which was significantly higher ($P = 0.006$) in SCC (Fig. 2). At the nuclear level, SCC showed significantly higher ($P = 0.016$) expression of bFGF than did adenocarcinoma, whereas the latter showed significantly higher expression of both receptors (FGFR1, $P < 0.0001$; FGFR2, $P = 0.0007$; Fig. 2).

One striking association observed was the correlation between the immunohistochemical expression of bFGF and the receptors with sex and smoking history in patients with adenocarcinoma (Table 2). Expression of nuclear bFGF and nuclear FGFR1 were significantly higher ($P = 0.03$ and $P = 0.02$, respectively) in women than in men. Although cytoplasmic FGFR1 expression was significantly higher ($P = 0.002$) in ever smokers than in never smokers, expression of nuclear FGFR1 ($P = 0.0001$) and nuclear FGFR2 ($P = 0.003$) was significantly higher in never smokers than in ever smokers. No significant associations between all three markers and gender or smoking status were found in patients with squamous carcinoma. For both NSCLC types, no correlation was detected between marker expression and age or TNM pathologic stage. In lung adenocarcinoma, FGFR1 nuclear expression positively and significantly (Spearman's correlation coefficient $r = 0.29$;
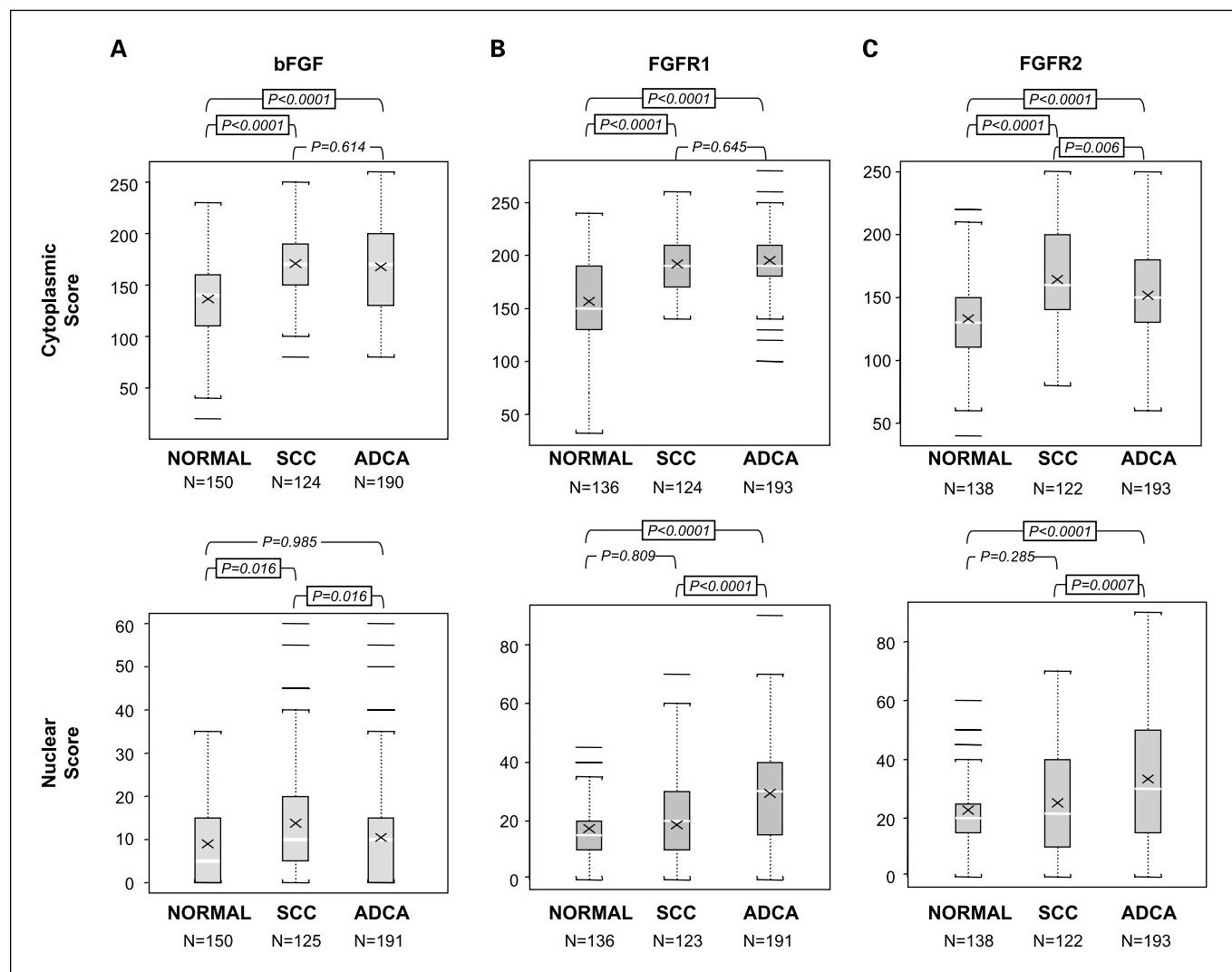
Fig. 2. Cytoplasmic (*top*) and nuclear (*bottom*) scores of immunohistochemical expression of bFGF (*A*), FGFR1 (*B*), and FGFR2 (*C*) in normal bronchial epithelia (*NORMAL*) obtained from lung cancer patients, SCC, and adenocarcinoma (*ADCA*) of the lung. The number of samples is indicated for each histologic group and marker. *P* values comparing normal epithelial and tumor histologic types are shown for all comparisons. Boxes around *P* values indicate statistical significance. *Bars,* 95% CI.

$P < 0.0001$) correlated with the presence of bronchioloalveolar subtype. Of interest, this correlation was independent of gender and remained significant ($P < 0.003$) after adjusting by smoking history in multivariate analysis (data not shown).

We analyzed the prognostic effects of the markers on disease outcomes using both a continuous score and a dichotomized score of the marker expressions using the cutoffs identified by Classification and Regression Tree algorithm. In patients with adenocarcinoma, both scores of cytoplasmic FGFR1 show a trend, although nonsignificant, effect on OS ($P = 0.070$ and $P = 0.068$, respectively, for the continuous and binary score, >195) and RFS ($P = 0.087$ and $P = 0.12$, respectively, for the continuous and binary measure) in the univariate setting. The Kaplan-Meier curves for the binary score shown in Fig. 4A indicate that high cytoplasmic FGFR1 was associated with a worse OS and a worse RFS. After adjusting for age, gender, smoking history, and pathologic stage, the effects of the cytoplasmic overexpression of FGFR1 on OS and on RFS remained marginally significant. Patients with high cytoplasmic FGFR1 had a worse OS [$P = 0.07$; hazard ratio (HR), 1.51; 95% confidence interval (95% CI), 0.97-2.33] and a worse RFS ($P = 0.05$; HR, 1.63; 95% CI, 1.00-2.67). Of interest in these patients, when treating as a continuous variable, each one-unit increase in the FGFR1 cytoplasmic score conferred a nonsignificant trend to a better OS ($P = 0.071$; HR, 1.01; 95% CI, 1.00-1.01) and a significantly worse RFS ($P = 0.02$; HR, 1.01; 95% CI, 1.00-1.02).

In patients with SCCs, we detected a more complex pattern of prognostic association. In the univariate analyses, nuclear FGFR1 (score > 17.5) and nuclear FGFR2 (score > 55) showed close to significant or significant effects on OS ($P = 0.061$ and 0.031 for nuclear FGFR1 and FGFR2, respectively). Nuclear FGFR1 and FGFR2 had a significant effect on RFS ($P = 0.027$ and 0.021, respectively). As the Kaplan-Meier curves for the binary score in Fig. 4B show, high nuclear expression of FGFFR1 was associated with a worse OS and a worse RFS. When age, gender, smoking history, and pathologic stage are compared, the nuclear overexpression of FGFR1 (score > 17.5) and FGFR2 (score > 55) significantly correlated with worse outcome in RFS and OS, respectively. The nuclear overexpression of FGFR1 conferred to patients a worse RFS ($P = 0.04$; HR, 1.97; 95% CI, 1.04-3.72), and the nuclear overexpression of FGFR2 correlated with worse OS ($P = 0.02$; HR, 2.54; 95% CI, 1.18-5.47) and RFS ($P = 0.02$; HR, 2.84; 95% CI, 1.18-6.86). In contrast, the cytoplasmic overexpression of bFGF (score > 175) and FGFR2 (score > 155) significantly correlated with better OS (bFGF: $P = 0.02$; HR, 0.55; 95% CI, 0.33-0.92 and FGFR2: $P = 0.008$; HR, 0.51; 95% CI, 0.31-0.83). When they are analyzed as a continuous variable, they also show trend, although nonsignificant, effects on OS and RFS. Again, high cytoplasmic bFGF is associated with a better OS ($P = 0.068$; HR, 0.99; 95% CI, 0.99-1.00 per unit increment). High cytoplasmic FGFR2 is associated with a better OS ($P = 0.086$; HR, 0.99; 95% CI, 0.99-1.00 per unit increment).

Correlations of markers' expression. In tumor specimens, we analyzed the correlation between the expression of all markers at cytoplasmic and nuclear localizations. Comparing all three markers, we detected a complex pattern of correlations of immunohistochemical expression in NSCLC that differed between the adenocarcinoma and SCC histologic types. A positive significant correlation between the cytoplasmic and nuclear overexpression of bFGF was detected in adenocarcinoma and of FGFR2 in both adenocarcinoma and SCC. In adenocarcinoma specimens, the correlation between bFGF and FGFR1 and between FGFR1 and FGFR2 was observed at cytoplasmic and nuclear localizations, whereas these correlations were found only at cytoplasmic localizations in specimens of SCC.

Discussion

As with some other angiogenesis pathways, the bFGF pathway has been shown previously to be activated in lung cancer (11–18). Although several reports showed high levels of
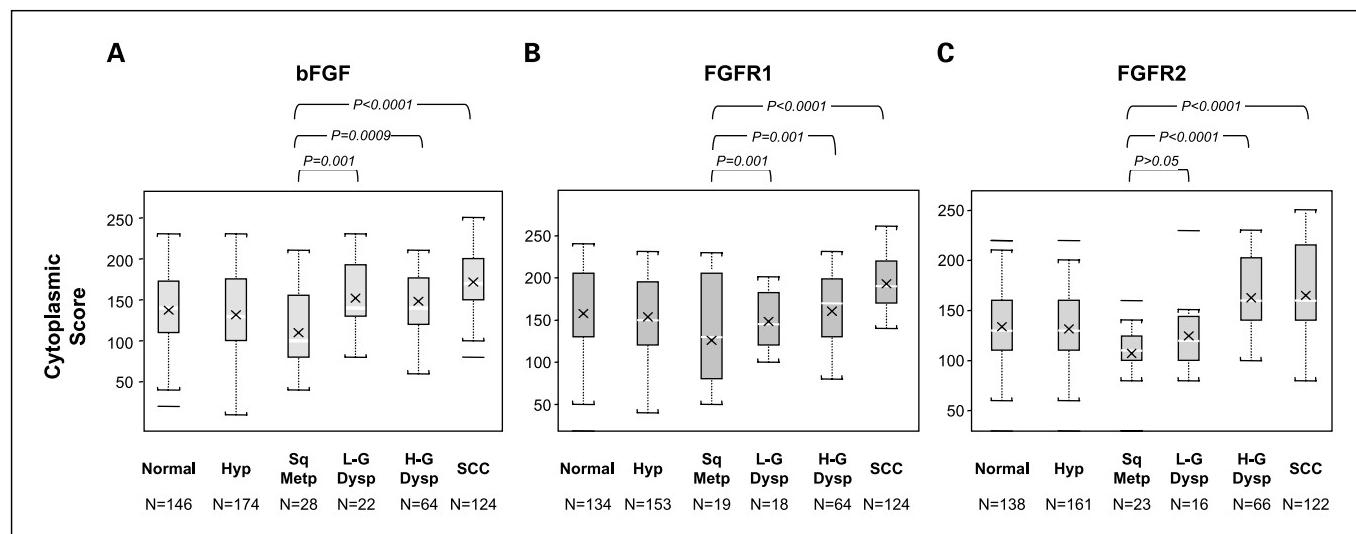
Fig. 3. Scores for cytoplasmic immunohistochemical expression of bFGF (*A*), FGFR1 (*B*), and FGFR2 (*C*) in bronchial respiratory epithelial lesions related to the pathogenesis of SCC of the lung: normal epithelium (*Normal*), hyperplasia (*Hyp*), squamous metaplasia (*Sq Metp*), low-grade dysplasia (*L-G Dysp*), high-grade dysplasia (*H-G Dysp*), and SCC. The number of samples is indicated for each histologic group and marker. Only significant *P* values for comparisons between squamous metaplastic and squamous dysplastic lesions and SCC are shown.

Table 2. bFGF, FGFR1, and FGFR2 immunohistochemical expression in NSCLC types by histology, sex, and smoking status

| | No. | Immunohistochemical markers, mean (SD) | | | | | |
| | | bFGF | | FGFR1 | | FGFR2 | |
		Cytoplasmic expression score*	Nuclear expression score †	Cytoplasmic expression score*	Nuclear expression score †	Cytoplasmic expression score*	Nuclear expression score †
Adenocarcinoma							
Sex							
Male	74	169.6 (47.6)	8.1 (9.4)	196.1 (27.7)	25.1 (18.9)	151.4 (37.4)	32.4 (23.1)
Female	122	166.1 (46.3)	11.9 (12.7)	194.6 (28.6)	31.8 (19.7)	151.9 (35.0)	33.9 (23.0)
P		0.62	0.03	0.73	0.02	0.92	0.65
Smoking status							
Never	50	162.5 (45.0)	11.6 (9.6)	184.7 (24.2)	38.4 (17.6)	148.4 (33.6)	41.5 (24.1)
Ever	146	169.1 (47.3)	10.1 (12.3)	198.8 (28.6)	26.2 (19.4)	152.9 (36.6)	30.5 (22.1)
P		0.40	0.44	0.002	0.0001	0.45	0.003
SCC							
Sex							
Male	79	174.5 (33.1)	12.9 (13.2)	193.3 (24.9)	18.3 (14.2)	166.1 (36.0)	24.2 (17.3)
Female	46	164.3 (37.0)	15.1 (12.8)	189.1 (26.5)	19.0 (13.5)	160.9 (36.6)	26.9 (18.8)
P		0.12	0.36	0.36	0.80	0.44	0.43
Smoking status							
Never	6	176.7 (46.3)	5.8 (13.3)	185.0 (15.2)	17.5 (11.7)	141.7 (43.1)	12.3 (9.9)
Ever	118	170.2 (35.1)	14.1 (13.2)	192.3 (25.5)	18.5 (14.1)	165.6 (35.6)	25.8 (18.0)
P		0.66	0.13	0.48	0.86	0.11	0.07

*Cytoplasmic immunostaining expression was quantified using a four-value intensity score (0, 1+, 2+, and 3+) and the percentage (0-100%) of the extent of reactivity. The cytoplasmic expression score was obtained by multiplying both intensity and reactivity extension values (range, 0-300).
† Nuclear immunostaining expression was quantified using a range of 0 to 100 according to the percentage of positive nuclei.

expression of bFGF and FGFR1 in NSCLC, the precise role of these signaling molecules in the pathogenesis and progression of this tumor is still unclear (12–18). Eight studies that conducted immunohistochemical expression analyses of NSCLCs (including adenocarcinomas and SCCs) described frequent bFGF expression in lung tumors, ranging from 49% to 77% of lung tumors (12–17, 21, 22). In four studies about FGFR1 immunohistochemical expression in NSCLC tissue specimens, these frequencies ranged from 50% to 81% (12–15). For FGFR2, only one immunohistochemical study was done involving 61 NSCLC specimens, 50% of which showed receptor expression (21). All of these studies used different arbitrary cutoffs to assess positive immunohistochemical expression; thus, the data are difficult to interpret.

In this study, we described high levels of immunohistochemical expression of bFGF, FGFR1, and FGFR2 in a large series of NSCLC specimens, including the two most frequent histologic types, SCC and adenocarcinoma, similarly to previous studies (12–17, 21, 22). To our knowledge, this is the first comprehensive analysis of bFGF, FGFR1, and FGFR2 in the same set of NSCLC specimens. Overall, we found higher levels of bFGF and FGFRs expression in tumor cells than in adjacent normal bronchial epithelia at cytoplasmic localization in both SCC and adenocarcinoma. These findings are consistent with the postulated mitogenic and angiogenic effects of the bFGF signaling pathway activation in human tumors, including NSCLC (8).

Lung cancers are believed to develop through a series of progressive pathologic changes (preneoplastic or precursor lesions) in the respiratory mucosa that result in an accumula-

tion of genetic abnormalities (23). These abnormalities are frequently extensive and multifocal throughout the respiratory epithelium, indicating a field effect or field cancerization phenomenon (23). Although sequential preneoplastic changes have been defined for centrally arising squamous carcinomas, they have been poorly documented for lung adenocarcinomas (4). Mucosal changes in the large airways that may precede invasive SCC include squamous dysplasia and carcinoma *in situ* in the central bronchial airway (24, 25). In lung cancer pathogenesis, the genetic changes commence in histologically normal epithelium and are present in a similar frequency in mildly abnormal epithelia, including basal cell hyperplasia and squamous metaplasia (23). In our study, we describe, to our knowledge for the first time, a widespread expression of bFGF, FGFR1, and FGFR2 adjacent to histologically normal, mildly abnormal, and preneoplastic bronchial respiratory epithelium in NSCLC specimens. Our findings of a significant increase in the immunohistochemical expression of bFGF and FGFR proteins at cytoplasmic (bFGF, FGFR1, and FGFR2) and nuclear (FGFR2) localizations in respiratory epithelium with low-grade (cytoplasmic bFGF) and high-grade (cytoplasmic bFGF and FGFRs and nuclear FGFR2) squamous dysplastic, compared with metaplastic, bronchial epithelia, suggest that the activation of the bFGF pathway is an early event in the pathogenesis of SCC of the lung and could play a role in the angiogenic switch characteristic of tumor promotion.

Whereas most of the previous studies reported either cytoplasmic protein expression only (12, 13, 16) or did not indicate immunostaining cell localization (14, 15, 17, 22), we described the expression of bFGF, FGFR1, and FGFR2 at both

cytoplasmic and nuclear localizations in our NSCLC specimens. These findings are consistent with previous *in vitro* studies showing that bFGF and FGFR1 translocate to the cell nucleus (26). In addition, the nuclear localization of bFGF in cells has been known for many years (27, 28), and there is some evidence that this translocation is required for the induction of cell proliferation (29). More recently, it has been shown that on cell stimulation with bFGF, FGFR1, a plasma membrane-associated protein, is undergoing endocytosis to the cytosol and translocates to the cell nucleus along with its ligand bFGF (26). Within the nucleus, FGFR1 serves as a general transcriptional regulator that activates structurally distinct genes located on different chromosomes and stimulates multigene programs for cell growth and differentiation (26).

In the correlation analysis of the immunohistochemical expression of bFGF and its receptors and patients' clinicopathologic characteristics, we found, somewhat surprisingly, a significantly higher level of nuclear bFGF and FGFR1 expression in tumor specimens obtained from female patients than in specimens from male patients with adenocarcinoma. These findings link the phenomenon of angiogenesis with sex-related oncogenic mechanisms, including sex steroid hormones. In women, the effects of estrogen and progesterone in angiogenesis

pathways such as those associated with vascular endothelial growth factor have been established in normal endometrial and breast tissues (30–33) and are being investigated in the corresponding tumors (34, 35). In the bFGF pathway, much less is known about whether sex steroids are involved with the regulation of normal and malignant tissues. In the normal endometrium, bFGF mRNA has shown a cyclic variation, suggesting that at least mRNA levels may be regulated by sex steroids (36). In breast cancer, bFGF levels seem to be up-regulated compared with normal adjacent tissue and associated with a high expression of the estrogen receptor, suggesting a correlation between them (35). The role of hormones, particularly estrogen, as a risk factor for the development of lung cancer among women is an area of vigorous investigation (37), and in NSCLC, there is evidence of cross-talk between the estrogen receptor and other growth factor signaling pathways, including the epidermal growth factor receptor (38).

In addition, in adenocarcinoma specimens, we detected differences in the expression of the three markers and patients' smoking status, with cytoplasmic FGFR1 expression being significantly higher in smokers and nuclear FGFR1 and FGFR2 significantly higher in never smokers. To our knowledge, these associations have not been reported
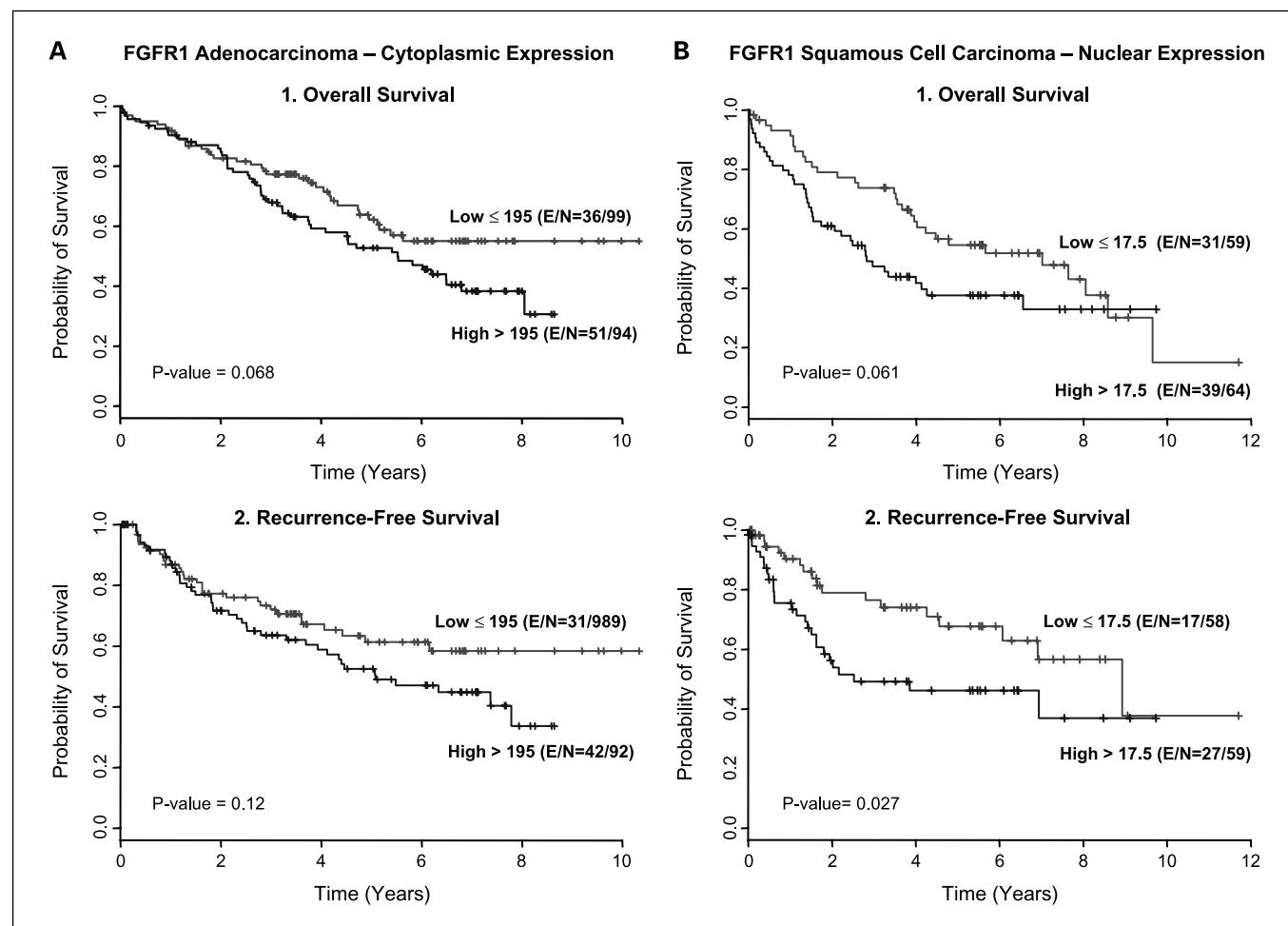

Fig. 4. Kaplan-Meier curves illustrating FGFR1 cytoplasmic protein expression for adenocarcinoma (*A*) and FGFR1 nuclear expression for SCC (*B*) patients.

previously. These differences highlight the potential differential role of these proteins in the pathogenesis of both smoking and non-smoking-related lung cancers. There is some evidence that links the bFGF pathway with smoking in lung diseases. It has been suggested that the bFGF pathway plays a role in regulating airway wall remodeling, especially in individuals with smoking-induced chronic obstructive peripheral disease, which is related to inflammation of the small airway (39). Of interest, Kranenburg et al. (40) conducted an immunohistochemical analysis of bronchial specimens obtained from patients with chronic obstructive peripheral disease and showed that FGFR1 immunolocalized in the cytoplasm of bronchial epithelial cells and in other cell components of the bronchial wall. We hypothesize that the inflammatory phenomenon occurring in smoking-damaged airway may be associated with the predominantly cytoplasmic localization of these receptors. Our finding of the differential expression of cytoplasmic FGFR2 and of both receptors at the nuclear level with smoking status in tumor specimens needs further study in epithelial lesions.

The prognostic effect of the expression of bFGF, FGFR1, and FGFR2 proteins in lung cancer has been studied previously in a limited number of tumor cases (14, 16, 17, 21). In NSCLC, high levels of bFGF expression, determined by ELISA in tissue extracts (n = 71; ref. 17) and by immunohistochemical expression in the cytoplasm of tumor cells (n = 119; ref. 16), correlated with worse OS. Similarly, high levels of FGFR1 and FGFR2 expression correlated with shorter OS in patients with NSCLC (n = 206; ref. 14) and lung adenocarcinomas (n = 30; ref. 21). When we investigated the correlation of the expression of these markers with OS and RFS, our findings suggested that the overexpression of FGFR1 in the cytoplasm of tumor cells correlated with worse OS only in patients with adenocarcinoma, similar to findings reported by Yamayoshi et al. (21). In contrast, for SCC, we found that the nuclear overexpression of FGFR1 and FGFR2 correlated significantly with worse outcome in both OS and RFS. Somewhat surprisingly, the cytoplasmic overexpression of bFGF and FGFR2 in squamous tumor cells correlated with better OS.

Although it seems paradoxical that a tumor that expresses high levels of bFGF and FGFR exhibits a better prognosis than a tumor that expresses low levels, we hypothesize that high levels of cytoplasmic expression may prevent nuclear localization of the proteins. Our data confirmed the notion that adenocarcinoma and SCC are not a homogeneous group of tumors and that they should always be examined separately for molecular targets and outcome.

We identified multiple correlations between the immunohistochemical expression of bFGF, FGFR1, and FGFR2 and their cytoplasmic and nuclear tumor cell localizations. All correlations but one were positive. Of interest, we detected a correlation between bFGF and FGFR1 expression in both NSCLC histologic types at the cytoplasmic level. These findings suggest that the cytoplasmic overexpression of the ligand bFGF in tumor cells may influence the expression and cell internalization of FGFR1. In lung adenocarcinomas, this association was also found at the nuclear level, emphasizing the role of bFGF in FGFR1 translocation to the cell nucleus (26). A similar pattern of association was detected between both FGFRs, a phenomenon that could be associated with preferential receptors dimerization and subsequent cytoplasmic internalization and nuclear translocation.

In summary, our findings indicate that bFGF, FGFR1, and FGFR2 are frequently overexpressed in NSCLC, although different patterns of expression are detected in its two major types. Our findings further suggest that bFGF signaling pathway activation is an early phenomenon in the pathogenesis of SCC of the lung. In addition, the frequent and early overexpression of bFGF and FGFR markers in patients with NSCLC suggests that the activation of the bFGF pathway, which has been proposed to facilitate the development of resistance to antiangiogenic therapy targeting the vascular endothelial growth factor pathway (41), is an attractive novel target for lung cancer therapeutic and chemopreventive strategies.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2006. CA Cancer J Clin 2006;56:106–30.

2. Minna JD, Gazdar A. Focus on lung cancer. Cancer Cell 2002;1:49–52.

3. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Tumours of the lung. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, editors. Pathology and genetics: tumours of the lung, pleura, thymus and heart. Lyon: IARC; 2004. p. 9–124.

4. Wistuba I. Genetics of preneoplasia: lessons from lung cancer. Curr Mol Med 2007;7:3–14.

5. Gazdar AF, Minna JD. Angiogenesis and the multistage development of lung cancers. Clin Cancer Res 2000;6:1611–2.

6. Keith RL, Miller YE, Gemmill RM, et al. Angiogenic squamous dysplasia in bronchi of individuals at high risk for lung cancer. Clin Cancer Res 2000;6:1616–25.

7. Merrick DT, Haney J, Petrunich S, et al. Overexpression of vascular endothelial growth factor and its receptors in bronchial dypslasia demonstrated by quantitative RT-PCR analysis. Lung Cancer 2005;48:31–45.

8. Dailey L, Ambrosetti D, Mansukhani A, Basilico C. Mechanisms underlying differential responses to FGF signaling. Cytokine Growth Factor Rev 2005;16:233–47.

9. Ribatti D,Vacca A, Rusnati M, Presta M. The discovery of basic fibroblast growth factor/fibroblast growth factor-2 and its role in haematological malignancies. Cytokine Growth Factor Rev 2007;18:327–34.

10. Mohammadi M, Olsen SK, Ibrahimi OA. Structural basis for fibroblast growth factor receptor activation. Cytokine Growth Factor Rev 2005;16:107–37.

11. Kuhn H, Kopff C, Konrad J, Riedel A, Gessner C, Wirtz H. Influence of basic fibroblast growth factor on the proliferation of non-small cell lung cancer cell lines. Lung Cancer 2004;44:167–74.

12. Takanami I,Tanaka F, HashizumeT, et al.The basic fibroblast growth factor and its receptor in pulmonary adenocarcinomas: an investigation of their expression as prognostic markers. EurJ Cancer 1996;32A:1504–9.

13. Takanami I, Tanaka F, Hashizume T, Kodaira S. Tumor angiogenesis in pulmonary adenocarcinomas: relationship with basic fibroblast growth factor, its receptor, and survival. Neoplasm 1997;44:295–8.

14. Volm M, Koomagi R, Mattern J, Stammler G. Prognostic value of basic fibroblast growth factor and its receptor (FGFR-1) in patients with non-small cell lung carcinomas. Eur J Cancer 1997;33:691–3.

15. Guddo F, Fontanini G, Reina C, Vignola AM, Angeletti A, Bonsignore G. The expression of basic fibroblast growth factor (bFGF) in tumor-associated stromal cells and vessels is inversely correlated with non-small cell lung cancer progression. Hum Pathol 1999;30:788–94.

16. Shou Y, HiranoT, Gong Y, et al. Influence of angiogenetic factors and matrix metalloproteinases upon tumour progression in non-small-cell lung cancer. Br J Cancer 2001;85:1706–12.

17. Iwasaki A, Kuwahara M, Yoshinaga Y, Shirakusa T. Basic fibroblast growth factor (bFGF) and vascular endothelial growth factor (VEGF) levels, as prognostic indicators in NSCLC. Eur J Cardiothorac Surg 2004; 25:443–9.

18. Bremnes RM, Camps C, Sirera R. Angiogenesis in non-small cell lung cancer: the prognostic impact of neoangiogenesis and the cytokines VEGF and bFGF in tumours and blood. Lung Cancer 2006;51:143–58.

19. BergerW, Setinek U, MohrT, et al. Evidence for a role

of FGF-2 and FGF receptors in the proliferation of non-small cell lung cancer cells. Int J Cancer 1999; 83:415–23.

20. Mountain CF. Revisions in the international system for staging lung cancer. Chest 1997;111:1710–7.

21. Yamayoshi T, Nagayasu T, Matsumoto K, Abo T, Hishikawa Y, Koji T. Expression of keratinocyte growth factor/fibroblast growth factor-7 and its receptor in human lung cancer: correlation with tumour proliferative activity and patient prognosis. J Pathol 2004;204:110–8.

22. Ito H, Oshita F, Kameda Y, et al. Expression of vascular endothelial growth factor and basic fibroblast growth factor in small adenocarcinomas. Oncol Rep 2002;9:119–23.

23. Wistuba II. Genetics of preneoplasia: lessons from lung cancer. Curr Mol Med 2007;7:3–14.

24. Colby TV, Wistuba II, Gazdar A. Precursors to pulmonary neoplasia. Adv Anat Pathol 1998;5:205–15.

25. Kerr KM. Pulmonary preinvasive neoplasia. J Clin Pathol 2001;54:257–71.

26. Stachowiak MK, Maher PA, Stachowiak EK. Integrative nuclear signaling in cell development—a role for FGF receptor-1. DNA Cell Biol 2007;26:811–26.

27. Friesel R, Maciag T. Internalization and degradation of heparin binding growth factor-I by endothelial cells. Biochem Biophys Res Commun 1988;151:957–64.

28. Walicke PA, Baird A. Internalization and processing of basic fibroblast growth factor by neurons and astrocytes. J Neurosci 1991;11:2249–58.

29. Wiedlocha A, Falnes PO, Madshus IH, Sandvig K, Olsnes S. Dual mode of signal transduction by externally added acidic fibroblast growth factor. Cell 1994; 76:1039–51.

30. Hyder SM, Murthy L, Stancel GM. Progestin regulation of vascular endothelial growth factor in human breast cancer cells. Cancer Res 1998;58:392–5.

31. Fujimoto J, Sakaguchi H, Hirose R, Ichigo S, Tamaya T. Progestins suppress estrogen-induced expression of vascular endothelial growth factor (VEGF) subtypes in uterine endometrial cancer cells. Cancer Lett 1999;141:63–71.

32. Hyder SM, Huang JC, Nawaz Z, et al. Regulation of vascular endothelial growth factor expression by estrogens and progestins. Environ Health Perspect 2000;108:785–90.

33. Welter H, Wollenhaupt K, Einspanier R. Developmental and hormonal regulated gene expression of fibroblast growth factor 2 (FGF-2) and its receptors in porcine endometrium. J Steroid Biochem Mol Biol 2004;88:295–304.

34. Fujimoto J, Toyoki H, Jahan I, et al. Sex steroid-dependent angiogenesis in uterine endometrial cancers. J Steroid Biochem Mol Biol 2005;93:161–5.

35. Smith K, Fox SB, Whitehouse R, et al. Upregulation of basic fibroblast growth factor in breast carcinoma and its relationship to vascular density, oestrogen receptor, epidermal growth factor receptor and survival. Ann Oncol 1999;10:707–13.

36. Nakamura J, Lu Q, Aberdeen G, Albrecht E, Brodie A. The effect of estrogen on aromatase and vascular endothelial growth factor messenger ribonucleic acid in the normal nonhuman primate mammary gland. J Clin Endocrinol Metab 1999;84:1432–7.

37. Stabile LP, Siegfried JM. Estrogen receptor pathways in lung cancer. Curr Oncol Rep 2004;6:259–67.

38. Stabile LP, Lyker JS, Gubish CT, Zhang W, Grandis JR, Siegfried JM. Combined targeting of the estrogen receptor and the epidermal growth factor receptor in non-small cell lung cancer shows enhanced antiproliferative effects. Cancer Res 2005;65:1459–70.

39. Hogg JC. Pathophysiology of airflow limitation in chronic obstructive pulmonary disease. Lancet 2004; 364:709–21.

40. Kranenburg AR, Willems-Widyastuti A, Mooi WJ, et al. Chronic obstructive pulmonary disease is associated with enhanced bronchial expression of FGF-1, FGF-2, and FGFR-1. J Pathol 2005;206:28–38.

41. Jubb AM, Oats AJ, Holden S, Koeppen H. Predicting benefits from anti-angiogenic agents in malignancy. Nat Rev Cancer 2006;6:626–35.

ORIGINAL ARTICLE

MOLECULAR SIGNATURES OF METASTASIS IN HEAD AND NECK CANCER

Stefano Colella, PhD,[1,2]* **Kristy L. Richards, PhD, MD,**[1,3]† **Linda L. Bachinski, PhD,**[1]
Keith A. Baggerly, PhD,[4,5] **Spiridon Tsavachidis, MSc,**[4] **James C. Lang, PhD,**[6]
David E. Schuller, MD,[6] **Ralf Krahe, PhD**[1,2,5-8]

[1] Department of Cancer Genetics, The University of Texas M. D. Anderson Cancer Center,
Houston, Texas. E-mail: rkrahe@mdanderson.org
[2] Human Cancer Genetics Program, Ohio State University, Comprehensive Cancer Center, Columbus, Ohio
[3] Division of Cancer Medicine, The University of Texas M. D. Anderson Cancer Center, Houston, Texas
[4] Department of Bioinformatics and Computational Biology, The University of Texas M. D. Anderson Cancer
Center, Houston, Texas
[5] Graduate Program in Human and Molecular Genetics, The University of Texas at Houston Graduate
School in Biomedical Sciences, Houston, Texas
[6] Molecular Biology and Cancer Genetics Program, Ohio State University, Comprehensive Cancer Center,
Columbus, Ohio
[7] Department of Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer
Center, Houston, Texas
[8] Graduate Program in Genes and Development, The University of Texas at Houston Graduate School in
Biomedical Sciences, Houston, Texas

Abstract: *Background.* Metastases are the primary cause of cancer treatment failure and death, yet metastatic mechanisms remain incompletely understood.

Methods. We studied the molecular basis of head and neck cancer metastasis by transcriptionally profiling 70 samples from 27 patients—matching normal adjacent tissue, primary tumor, and cervical lymph node metastases.

Results. We identified tumor-associated expression signatures common to both primary tumors and metastases. Use of matching metastases revealed an additional 46 dysregulated genes associated solely with head and neck cancer metastasis. However, despite being metastasis-specific in our sample set, these 46 genes are concordant with genes previously discovered in primary tumors that metastasized.

Conclusions. Although our data and related studies show that most of the metastatic potential appears to be inherent to the primary tumor, they are also consistent with the notion that a limited

For most solid tumors, tumorigenesis occurs as a multistep process including both genetic and epigenetic alterations, which together provide cancer cells with a selective growth advantage. Transformation of normal cells into malignant cells is characterized by a limited set of acquired capabilities, which are likely shared among different tumor types.[1] However, metastases to distant organ sites are ultimately responsible for most cancer deaths.[2] Despite much progress, important questions about metastatic progression remain to be answered.[2–4] With respect to metastatic origins, 2 conflicting views have been proposed. The "rare metastatic variants" model holds that rare metastatic cells preexist within the primary tumor and subsequently progress to metastases by the process of selection.[5,6] This model has recently been challenged by the results of global gene expression profiling in various cancer types. These studies have provided evidence for the independent evolution of 2 classes of primary tumors, those that are predisposed to metastasize and those that are not ("primary tumor predisposition" model).[7–11] We have attempted to address the mechanism(s) of metastasis using expression profiling of genetically matched primary tumors, lymph node metastases, and adjacent normal tissue.

In this study, we focused on squamous cell carcinoma (SCC) of the head and neck (HNSCC) as a prototypic epithelial solid tumor.[12–14] Several molecular alterations associated with HNSCC have been identified, but a complete understanding of the complex molecular events underlying tumorigenesis and metastasis is still lacking.[13,15,16] Cervical lymph node metastases are the single most adverse independent prognostic indicator for local recurrence and/or distant metastasis.[17] A reliable method to identify those tumors most likely to recur or to produce distant metastases would significantly enhance the choice of effective treatment options and, therefore, outcomes for the patients.

Several groups have attempted to predict lymph node metastases retrospectively by identifying a metastatic signature in primary tumor tissues. However, in the majority of these studies, the predictive signatures have not been compared with the expression signature that is actually present in the metastases themselves. To identify the genetic changes underlying HNSCC tumor initiation, progression, and metastasis directly, we generated comprehensive expression profiles for 70 samples, including primary tumors, genetically matched normal adjacent mucosae and genetically matched cervical lymph node metastases. Rigorous statistical analysis revealed unique expression profiles differentiating normal mucosae from primary tumor and metastatic lesions. Genetically matched primary tumor/metastasis samples (ie, from the same patients) showed highly correlated individual profiles, consistent with the notion that the metastatic potential is already encoded in the primary tumor. However, we also identified 46 genes whose dysregulated expression was specific for head and neck cancer metastasis, suggesting that the acquisition of an additional limited number of clonal changes results in the final metastatic cell(s). Findings are compared with those of other published studies, and the implications for the metastatic process in solid tumors and the utility for clinical prognostication based on gene expression in primary tumors are discussed.

PATIENTS AND METHODS

Patient Biopsy Samples. Informed consent was obtained from all patients prior to sample collection in accordance with the guidelines set by the Institutional Review Board of The Ohio State University (OSU). Biopsy samples from surgical resections were collected and banked through the OSU Head and Neck Cancer Tissue Bank. For each patient, the primary tumor, matching metastatic lymph node, and adjacent normal mucosa were collected at the same time and processed concurrently to minimize confounding clinical or technical variability. Primary tumor tissue and tissue from metastatic lymph nodes were subject to examination by a board-certified pathologist to confirm tissue histology and cellular representation. Clinically appearing tumor tissue was taken distant from the tumor margin and was macro-dissected to provide tissue that contains >70% tumor cells. Normal tissue was clinically appearing normal adjacent tissue taken at least 3 cm from the tumor margin, usually contralateral. All samples were immediately placed on ice and, after removal of portions needed for pathological diagnosis, were snap-frozen in liquid nitrogen within 30 minutes of devascularization. Snap-frozen samples were held at $-80°C$ for long-term storage. We used samples from 5 representative anatomical locations (8 larynx, 8 tongue, 6 pharynx, 5 tonsil, 1 palate) from 27

patients. For 1 patient, we had tumor samples before (N13, P13) and after treatment (N3, P3, M3). Supplementary Table S1 summarizes the clinical and demographic features of the patients.

Sample Preparation. In all, we profiled 70 samples from 27 patients—28 primary tumors, 28 matching normal adjacent mucosae, and 14 matching lymph node metastases. RNA was extracted from snap-frozen tissue samples using the TriZol Reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's suggestions. RNA was further purified using the RNeasy kit (Qiagen, Valencia, CA). For hybridization to the HuGeneFL and U95Av2 GeneChip arrays (Affymetrix, Santa Clara, CA) cRNA target preparation was according to the Affymetrix protocol. Thirty-five samples (12 normal adjacent mucosae, 12 primary tumors, and 11 metastatic lymph nodes from the same patients) were profiled on the HuGeneFL platform, and 35 samples (16 normal adjacent mucosae, 16 primary tumors, and 3 metastatic lymph nodes from the same patients) were profiled using the U95Av2 platform, interrogating approximately 6600 and 12,500 genes, respectively.

Data Analysis Across Platforms. All original GeneChip.cel files are available at Array Express (accession number: E-MEXP-44; www.ebi.ac.uk/arrayexpress/). The .cel files were processed and quantified with the dChip analysis package using the Li–Wong full model (Supplementary Data S1 and S2).[18,19] This choice was based on our previous finding that this method produced the most reliable results for the analysis of oligonucleotide arrays.[20] All statistical analyses were performed with MATLAB (MathWorks, Natick, MA). We analyzed the 2 datasets generated on the HuGeneFL and U95Av2 platforms using standard exploratory techniques, including principal component analysis (PCA), pairwise distance matrix (PDM), and hierarchical cluster (HC) analyses. To contrast the different sample types (normal adjacent tissue, primary tumor, and lymph node metastasis), we combined the 2 datasets to generate a list of genes shared between the 2 GeneChip platforms (Supplementary Data S3).

To combine data from multiple probe sets across both platforms, we used a "pseudo" probe set approach (Supplementary Data S4).[21] Briefly, the probe sequence information supplied by Affymetrix was used to identify individual probes with identical sequences present on both GeneChip platforms. Common probes were matched against the latest build of UniGene to assemble pseudo probe sets consisting of all common probes mapping to a given UniGene cluster. Because HuGeneFL and U95Av2 interrogate 20 and 16 probes per probe set/gene, respectively, the number of common probes in a pseudo probeset can vary. Therefore, computations were performed only on pseudo probe sets with at least 3 probes, interrogating a total of 4509 unique genes with 7871 pseudo probe sets. We used the robust multichip average (RMA) methodology[22] for quantification of pseudo probe set intensities. RMA enables correction of spatial effects on individual chips and matching of values across chip types using quantile normalization. After combining the data using this approach, quantifications for the 2 chip types are on the same scale.

To analyze combined results for primary tumor and normal adjacent tissue samples for the genes shared between array types, we used nonparametric tests, which do not require the gene quantification values to be the same across array types, including a sign test that counts the number of times the gene expression in primary tumor is higher or lower than the gene expression in the paired normal adjacent tissue sample. We focused on those genes for which the sign of the difference was the same in at least 25 of 28 pairs (90%). This cutoff is conservative in that the chance of a given gene being higher in 25/28 primary tumor samples is $1.3720\,e-05$. Adjusting for multiple comparisons, the number of differentially expressed genes to be expected by chance is 0.1, whereas we actually see several hundred genes. A parallel argument can be made for the genes found to be higher in normal adjacent tissue than primary tumor samples. The sign test exploits the paired nature of the data by using just the paired differences as inputs. In this respect, the test is more specific than PCA, which does not take into consideration the patient-to-patient heterogeneity. We also used approaches that did not exploit the paired nature of the samples, such as Wilcoxon rank sum tests used to compare normal adjacent tissue and primary tumor samples within the same array type and then combined the p values.

For the primary tumor versus lymph node metastasis comparison, we first used a sign test that exploited the paired nature of the data similar to the test used for the normal adjacent tissue versus primary tumor comparison. For comparisons within a chip type, we also applied signed rank tests to exploit the magnitude order. Applying the sign test, we first focused on genes found to be higher in metastases in at least 12 of 14 cases

(86%). With this cutoff, the chance of a given gene being higher in 12/14 lymph node metastasis samples is 6.4697 e−03. A Wilcoxon rank sum test that ignores the paired nature was also used to compare normal adjacent tissue versus lymph node metastasis samples.

RESULTS

Normal Adjacent Tissue is Distinct from Primary Tumor and Metastasis. Using nonparametric tests to identify differentially expressed genes and comparing the primary tumors with their genetically matched normal adjacent mucosae, we identified 414 probe sets, representing 345 unique genes, which showed significant gene expression differences in the same direction in at least 25 of 28 samples (p value = 1.372×10^{-5}) (Supplementary Table S2). We also applied the rank filter to the pseudo probe set data and identified 338 distinct Unigene clusters (genes). Of these 338 genes, 249 were also on the 345-gene list, using the Entrez Gene ID as the common identifier (Supplementary Table S2). Among these were several genes previously implicated in HNSCC carcinogenesis, including *EGFR*, *CCNB1*, and *STAT1*.[14,23–25]

Genetically Matched Primary Tumors and Metastases Share a Common Expression Profile. To determine the overall relationships between samples, we performed PCA (Figure 1A) and paired distance matrix (PDM, Figure 1B) analysis using all genes from the pseudo probe set approach (4509) and all 70 samples. PCA revealed specific gene signatures that distinguished primary tumors and metastases from normal adjacent tissue. PDM analysis showed more variation among the primary tumors and the metastases than among the genetically matched normal adjacent tissue. Yet, at the same time, high correlation between primary tumor and metastasis from the same patients was clearly apparent (reflected in 2 diagonal lines of blue squares comparing lymph node metastasis and primary tumor in Figure 1B). Two-way hierarchical cluster analysis based on the 338 tumor-specific genes obtained from the pseudo probe set approach confirmed the primary tumor/metastasis pairing for the majority of samples (9 of 14). Interestingly, 2 primary tumors extracted from the same patient before and after treatment (P13 and P3) clustered closely as well (Figure 2). When we performed hierarchical clustering by chip type using the 345 unique genes obtained using a rank

test, sample clustering was consistent with that obtained for all samples combined; 10 of 14 primary tumors were paired with their corresponding metastases (data not shown) indicating the robustness across platforms. In fact, this degree of clustering (10/14 or 71%) is nearly as high as that obtained by comparing 2 separate biopsies of the same primary tumor by expression profile clustering (80% to 90%).[26]

Identification of Metastasis-Specific Genes Through Paired Analysis. To identify genes specifically associated with the metastatic process, we compared the 14 available cervical lymph node metastases to their genetically matched primary tumors. Because of patient-to-patient heterogeneity in gene expression levels, the use of genetically unmatched primary tumors and metastases or retrospective analysis of primaries alone can be confounding. Figure 3 shows examples of variation in expression levels across individual samples for several representative genes. For each patient sample set, the general trend of dysregulation was the same. However, the range of actual levels for primary tumor and lymph node metastasis across all sets of sample overlaps, resulting in a loss of sensitivity if the paired nature of the data is ignored. To determine the power of a nonpaired relative to a paired analysis in identifying dysregulated genes, we evaluated 3 statistical tests—the Wilcoxon rank sum test, which uses a nonpaired approach, and the signed rank test and sign test, both of which take advantage of the paired nature of our sample set. Both of the latter tests, which utilize the paired design, perform statistically better than the Wilcoxon rank sum test, which ignores the paired design (data not shown).

To combine data for all 14 lymph node metastases generated on the 2 different platforms, we used the sign test to perform the paired analysis. We identified genes that showed statistically significant differences in at least 11 of 14 samples (p value = .0287) between primary tumors and matching metastases (Supplementary Table S3). Because many of the genes that were different by the sign test had absolute differences that were relatively small, an additional filter of a mean-fold change of ≥ 1.5 was applied. This resulted in a list of 46 metastasis genes—16 upregulated and 30 downregulated (Table 1). Ten of the 46 genes were represented by at least 2 probe sets. Several of these genes are known to be involved in the metastatic process of other tumor types, including *ENPP2* and *CXCR4*, both of which were
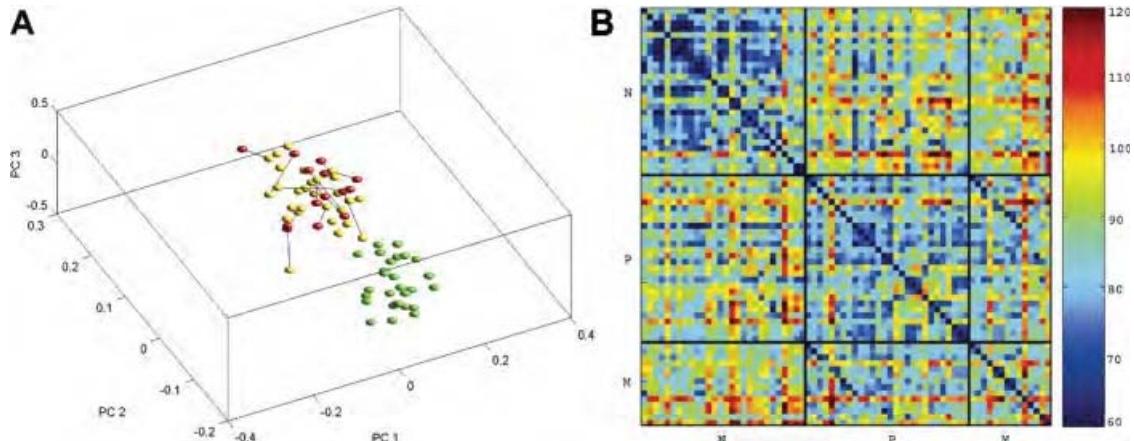
FIGURE 1. Analyses of expression profiles generated from the 4509 pseudo probe set genes. (**A**) Principal component analysis (PCA) showing all 70 samples plotted for the first 3 principal components derived using the expression matrix. The expression values for each gene were log-transformed and centered before computing the principal components. Green dots are normals, yellows are primary tumors, and reds are metastases; lines also join matched primary tumors and metastases. There is a clear separation between the genetically matching normal adjacent samples and tumor samples (primary tumors and metastases). Most metastases are closer to their corresponding primary tumors than to other metastasis profiles. (**B**) Pairwise distance matrix (PDM) between all 70 samples, using Euclidean distance applied to robust multichip average (RMA) quantifications of all 4509 pseudo probe set genes. The order of normals (N), primary tumors (P), and metastases (M) is the same within blocks. Distances between normals and tumors of all types are greater than distances between normals and other normals, or between tumors and other tumors. The comparison of primary tumors with metastases provided evidence for a tumor-specific signature indicated by shorter distances from primary tumors to their corresponding metastases compared with other metastases. This closer relatedness is visible as a darker blue diagonal in the primary tumor versus lymph node metastasis comparison in the upper left corner of the block of distances between primary tumor and lymph node metastasis.
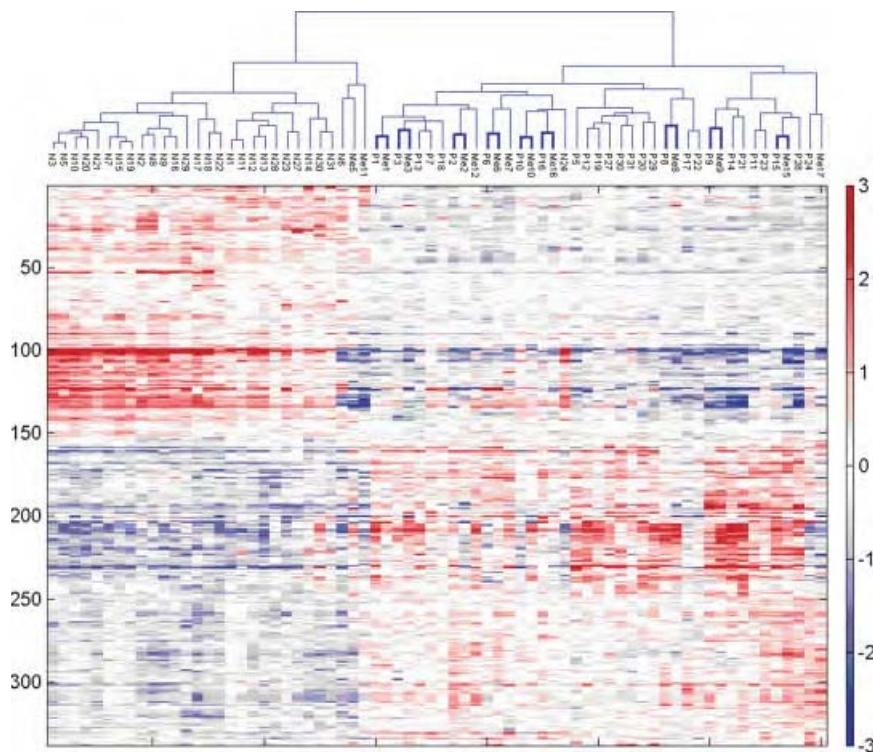
FIGURE 2. Two-way hierarchical cluster analysis of gene expression profiles generated on pseudo probe set gene expression data. Hierarchical clustering of all 70 samples (cell lines to the side) from both chip types using Euclidean distance and complete linkage applied to robust multichip average (RMA) quantifications of the 338 pseudo-probe sets passing the rank filter. Normals (N) cluster to the left, tumors to the right. Of particular interest is that most primary tumors (P) and lymph node metastases (M) from the same patient show pairing (9 of 14, 64% highlighted by thicker lines). Two primary tumor samples obtained from the same patient before and after treatment (P13 and P3) are also very near in the hierarchical samples cluster, suggesting a patient-specific tumor signature.
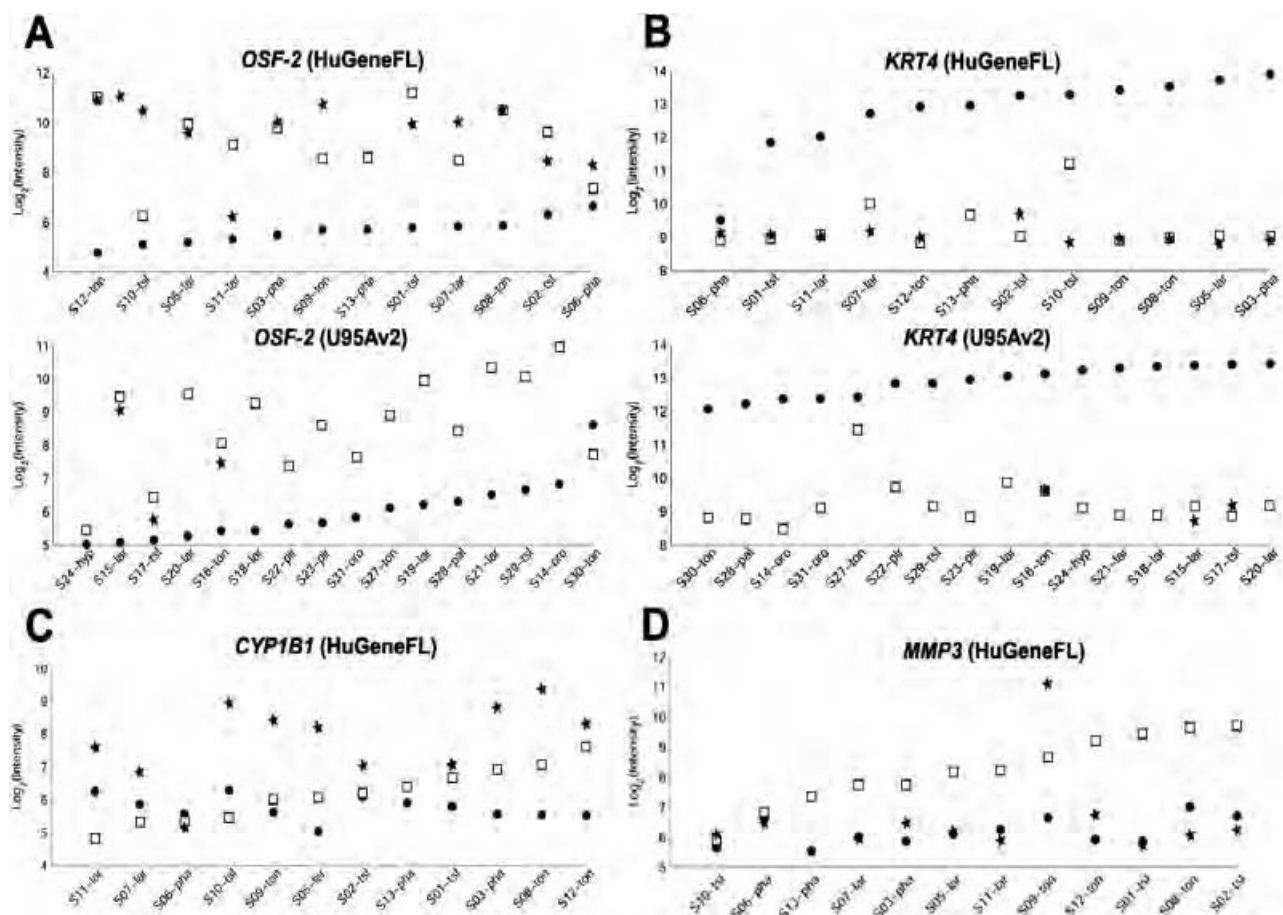
FIGURE 3. Expression levels for specific genes in sets of genetically matching normal adjacent mucosae (●), primary tumors (□), and metastatic lymph nodes (★). The data plotted for different sample sets (S) show considerable variation in expression levels across individual sets for the same gene in the same tissue types. For each patient sample set, the general trend of dysregulation (overexpression or underexpression) is in the same direction, whereas levels across all patient sets overlap. The plots highlight the utility of the matched-based analysis approach, which improves the sensitivity of the analysis. Representative genes (**A**) overexpressed (*OSF-2*) or (**B**) underexpressed (*KER4*) in primary tumor and lymph node metastasis relative to their genetically matched normal adjacent tissue sample, in ascending order of expression in the normal tissue. Representative genes (**C**) overexpressed (*CYP1B*) or (**D**) underexpressed (*MMP3*) in lymph node metastasis relative to their genetically matched primary tumor sample, in ascending order of expression in the primary tumor.

upregulated, and *IL24*, which was downregulated.[27–29] Interestingly, of the 46 metastasis-specific genes, 9 (20%) are muscle-related genes, the highest single dysregulated functional category. This is possibly associated with the increased locomotory and invasive phenotypes of metastatic cells.[30]

Comparison of Head and Neck Cancer Metastasis-Specific Genes with Other Published Head and Neck Cancer Metastasis Signatures. A variety of approaches have been undertaken to define a metastasis signature in HNSCC. Most groups have approached the question using primary tumors only, along with supervised analyses to identify genes that are differentially expressed in primary tumors associated with lymph node metastases,

versus those without lymph node metastases.[31–35] Another approach compared immortal with mortal cell lines established from HNSCC samples.[36] A third approach, similar to ours, compared matched primary tumors to metastases and identified a single differentially expressed gene, *MTA1*.[37] The specific details of these studies are compared in Supplementary Table S4. Comparing the resulting lists of differentially expressed genes from these reports with each other and with our 46-gene list yielded relatively few overlapping genes (Supplementary Figure S2). However, when cellular processes and pathways, as opposed to individual gene products, were compared between all of these studies, including ours, the results were strikingly similar. Using Ingenuity Pathway Analysis

Table 1. HNSCC metastasis genes.

Hu6800 probe ID	U95Av2 probe ID	Gene symbol	Title	Entrez gene	Location	Average positive fold change
16 Upregulated genes						
U03688_at	859_at	CYP1B1	Cytochrome P450, subfamily I (dioxin-inducible), polypeptide 1 (glaucoma 3, primary infantile)	1545	Chr:2p21	2.5564
U03688_at	40071_at					2.5166
L35594_at	41124_r_at	ENPP2	Ectonucleotide pyrophosphatase/phosphodiesterase 2 (autotaxin)	5168	Chr:8q24.1	1.6961
L35594_at	41123_s_at					1.6668
J03507_at	37394_at	C7	Complement component 7	730	Chr:5p13	2.1884
M12529_at	608_at	APOE	Apolipoprotein E	348	Chr:19q13.2	1.7824
X13839_at	32755_at	ACTA2	Actin, α 2, smooth muscle, aorta	59	Chr:10q23.3	1.6111
X64877_s_at	36341_s_at	HFL3	H factor (complement)-like 3	3080	Chr:1q31–q32.1	1.6062
Z30426_at	37645_at	CD69	CD69 antigen (p60, early T-cell activation antigen)	969	Chr:12p13–p12	1.5928
X57809_s_at	31459_i_at	IGL@	Immunoglobulin lambda locus	3535	Chr:22q11.1–q11.2	1.5676
X57809_s_at	31344_at					1.5276
U12535_at	1467_at	EPS8	Epidermal growth factor receptor pathway substrate 8	2059	Chr:12q23–q24	1.5052
U37546_s_at	1717_s_at	BIRC3	Baculoviral IAP repeat-containing 3	330	Chr:11q22	1.9534
U65404_at	137_at	KLF1	Kruppel-like factor 1 (erythroid)	10661	Chr:19p13.13–p13.12	1.9473
X53331_at	36683_at	MGP	Matrix Gla protein	4256	Chr:12p13.1–p12.3	1.8900
X89101_s_at	1441_s_at	TNFRSF6	Tumor necrosis factor receptor superfamily, member 6	355	Chr:10q24.1	1.6919
M98539_at	216_at	PTGDS	Prostaglandin D2 synthase 21 kDa (brain)	5730	Chr:9q34.2–q34.3	1.638
U10550_at	37279_at	GEM	GTP binding protein overexpressed in skeletal muscle	2669	Chr:8q13–q21	1.5689
L06797_s_at	649_s_at	CXCR4	Chemokine (C–X–C motif) receptor 4	7852	Chr:2q21	1.5148
30 Downregulated genes						
M20642_s_at	40157_s_at	MYL1	Myosin, light polypeptide 1, alkali; skeletal, fast	4632	Chr:2q33–q34	−4.2466
M20642_s_at	40158_r_at					−4.5890
V00594_s_at	39081_at	MT2A	metallothionein 2A	4502	Chr:16q13	−1.5118
V00594_at	39081_at					−1.5128
X90568_at	40795_at	TTN	Titin	7273	Chr:2q24.3	−3.2370
S73840_at	39101_at	MYH2	Myosin, heavy polypeptide 2, skeletal muscle, adult	4620	Chr:17p13.1	−2.9306
X66141_at	36640_at	MYL2	Myosin, light polypeptide 2, regulatory, cardiac, slow	4633	Chr:12q23–q24.3	−2.0088
M24069_at	39839_at	CSDA	Cold shock domain protein A	8531	Chr:12p13.1	−1.5549
M33772_s_at	41748_at	TNNC2	Troponin C2, fast	7125	Chr:20q12–q13.11	−1.5510
X53586_rna1_at	33411_g_at	ITGA6	Integrin, α 6	3655	Chr:2q31.1	−1.5346
X53586_rna1_at	33410_at					−1.5240
X64177_f_at	39594_f_at	MT1H	Metallothionein 1H	4496	Chr:16q13	−1.5339
U35637_s_at	38461_at	NEB	Nebulin	4703	Chr:2q22	−2.9512
X53961_at	37149_s_at	LTF	Lactotransferrin	4057	Chr:3q21–q23	−2.7714
X51441_at	33272_at	SAA1	Serum amyloid A1	6288	Chr:11p15.1	−2.5386
X51441_s_at	33272_at					−2.3024
X05232_at	437_at	MMP3	Matrix metalloproteinase 3 (stromelysin 1, progelatinase)	4314	Chr:11q22.3	−2.3736
M69225_at	40304_at	BPAG1	Bullous pemphigoid antigen 1, 230/240kDa	667	Chr:6p12–p11	−2.2511
M69225_at	32782_r_at					−2.0523
L20861_at	31862_at	WNT5A	Wingless-type MMTV integration site family, member 5A	7474	Chr:3p21–p14	−2.2151

Table 1. HNSCC metastasis genes (*Continued*).

Hu6800 probe ID	U95Av2 probe ID	Gene symbol	Title	Entrez gene	Location	Average positive fold change
30 Downregulated genes (contd.)						
L20861_at	1669_at		Actin, alpha, cardiac muscle	70	Chr:15q11–q14	−1.5605
J00073_at	39063_at	ACTC	Interleukin 8	3576	Chr:4q13–q21	−2.0703
Y00787_s_at	35372_r_at	IL8	S100 calcium binding protein A2	6273	Chr:1q21	−2.0136
Y07755_at	35726_at	S100A2	Ras-related associated with diabetes	6236	Chr:16q22	−2.0042
L24564_at	39528_at	RRAD	Calbindin 1, 28 kDa	793	Chr:8q21.3–q22.1	−1.9363
X06661_at	36570_at	CALB1	Collagen, type XVII, alpha 1	1308	Chr:10q24.3	−1.8359
M91669_s_at	41618_at	COL17A1	Neuroepithelial cell transforming gene 1	10276	Chr:10p15	−1.8089
U02081_at	33894_at	NET1	Tropomyosin 2 (beta)	7169	Chr:9p13.2–p13.1	−1.7926
X06825_at	32312_at	TPM2				−1.7862
X06825_at	32314_g_at					−1.7145
U16261_at	41848_f_at	IL24	Interleukin 24	11009	Chr:1q32	−1.7819
M93056_at	33305_at	SERPINB1	Serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 1	1992	Chr:6p25	−1.7538
M11433_at	38634_at	RBP1	Retinol binding protein 1, cellular	5947	Chr:3q23	−1.7342
X58377_at	35464_at	IL11	Interleukin 11	3589	Chr:19q13.3–q13.4	−1.5896
U41060_at	1798_at	SLC39A6	Solute carrier family 39 (zinc transporter), member 6	25800	Chr:18q12.1	−1.5377
M80244_at	32186_at	SLC7A5	Solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	8140	Chr:16q24.3	−1.5284
U89916_at	39579_at	CLDN10	Claudin 10	9071	Chr:13q31–q34	−1.5012

(Ingenuity Systems; www.ingenuity.com), genes involved in cellular processes including cancer, cell cycle, and cell-to-cell signaling were disproportionately overrepresented on gene lists from all 7 studies. Their degree of overrepresentation was generally similar, as indicated by the similarity in the significance level. Results of the comparison are shown in Supplementary Figure S1.

DISCUSSION

Although the primary risk factors for HNSCC (tobacco and alcohol consumption) are well recognized, and considerable progress has been made, understanding of the molecular events and mechanisms underlying the multistep carcinogenesis process in HNSCC is still incomplete and prognosis continues to be poor. The identification of genes and pathways involved in tumor initiation, progression and especially metastasis is therefore an important goal. Several groups have published distinct gene expression signatures associated with the presence of lymph node or distant metastases.[31–35,38] We profiled genetically matched uninvolved adjacent tissues, primary tumors and metastatic lymph nodes from a representative set of HNSCC patients. To our knowledge, this is the largest data set yet utilizing such genetically matched samples.[39–41]

Genes Associated with Tumor Progression. Genes that showed concordant levels of dysregulation in the primary tumors and metastases relative to their matched normal tissues are likely involved in tumor initiation and maintenance. Tumor progression in this sample set appeared to be particularly associated with the dysregulation of genes involved in cell adhesion, cell-cycle regulation, and remodeling of the extracellular matrix (ECM) and cytoskeleton. Some of the dysregulated genes and cellular pathways in HNSCC are shared with other solid tumors, suggesting that there are a limited number of cellular functions and pathways contributing directly to tumor initiation and progression. The identification of genes previously implicated in other cancers, and HNSCC specifically, confirms the validity of our approach. In addition, we identified several genes not previously associated with HNSCC tumorigenesis, some of which constitute potential HNSCC biomarkers.

Primary Tumors are Very Similar to Their Corresponding Lymph Node Metastasis. SCC of the head and neck can be considered a prototypic

solid tumor of epithelial origin that results from the accumulation of multiple genetic and epigenetic alterations.[15,42] As with many other solid tumors, the presence of cervical lymph node metastasis is considered the single most important indicator for local recurrence and/or distant metastasis.[17] The remarkable similarity that we observed between primary tumor and lymph node metastasis of the same patient, reflected in the paired clustering for the majority of samples, is consistent with the notion that the metastatic potential is already encoded in the bulk of the primary tumor. This finding tends to contradict the "rare metastatic variants" model and supports the "primary tumor predisposition" model.

Genes Associated with Metastasis. Despite the high similarity between primary tumors and matched metastases, incorporating genetically matched samples in our study design enabled us to overcome inter-patient tumor heterogeneity and identify 46 genes that were significantly differentially expressed in lymph node metastases of HNSCC patients. Some of the 46 genes, and the cellular pathways in which their protein products function, have been implicated in tumor progression and metastasis in previous studies. These include genes that individually have been correlated with metastatic behavior and organ-specific metastasis in other cancers of epithelial origin. Several of these genes (*ENPP2*, *CXCR4*, *LTF*, *S100A2*, and *IL24*) have been identified as mechanistically important in the cascade of events that drives metastasis, and many of the other genes on our list have been implicated in tumorigenesis.[43–47] However, most of them were previously unrecognized as genes contributing to the development of HNSCC metastasis.

Whether our newly identified metastasis-associated genes contribute functionally to metastatic progression, not only in HNSCC but in other cancers as well, will require further study. Nonetheless, our findings are encouraging, because they suggest the involvement of a limited set of genes and pathways that might be common therapeutic targets in different solid cancers. We were able to identify these genes because of our paired design, which enabled us to overcome inter-tumor heterogeneity, a confounding factor in the analysis of unmatched samples.

Comparison with Previously Published Studies. Given that several HNSCC metastatic signatures have now been published,[31–36] it appeared

worthwhile to examine the studies for overlap as a means to derive information that may have been missed in the analysis of any single study in isolation. These HNSCC metastatic signatures were derived using several different profiling platforms and analytical approaches (ie, supervised analysis of primary tumor signatures based on association with presence of lymph node metastases, comparison of immortal versus mortal HNSCC samples, and in our case, comparison of metastases with paired primary tumors). These diverse study designs can explain some of the discrepancies seen in the direction of gene expression changes and strengthen the significance of the genes that are independently discovered in multiple studies (Supplementary Figure S2). Similarities seen between all 7 studies when pathways, rather than individual gene products, were compared indicates that higher-level systems analysis will likely be essential for the understanding and development of clinical applications from primary gene expression data. Interestingly, the metastasis signatures for different solid tumors[10] and ours for HNSCC, showed a similar enrichment of muscle-related and metallothionein genes. The fact that tumor progression is reflected in the remodeling of the ECM is consistent with the observations that epithelial–mesenchymal interactions are critical factors of tumor cell behavior.[48–50] The theme that is emerging is one of loss of cell adhesion and acquisition of increased cell motility, which endows the cell with the ability to migrate, invade, and home to specific organs—in the case of HNSCC, first lung then bone.

Progression to Metastasis in HNSCC. Clearly, head and neck cancer tumorigenesis is a multistep process that results from the accumulation of multiple genetic and epigenetic alterations. Together, these events provide the tumor cell with a selective growth advantage, which drives the development of the primary tumor and metastasis. The striking similarity of the expression profiles between primary tumors and their genetically matched metastases, and the identification of the 46 metastasis-specific genes together support the hypothesis that, compared with the development of the primary tumor, fewer additional clonal changes are necessary to yield metastatic cells. Expression profiling of cancer cell lines selected in vivo for their metastatic potential also indicated that changes in only a limited number of genes and cellular functions underlie the acquisition of the metastatic phenotype.[44,51,52] Similarity

between our metastatic signature (derived from metastases themselves) and other metastatic signatures (derived from primary tumors alone) suggests that the order of these changes may be variable, because our study found these expression changes only in metastases, but others identified them as already present in the primary tumor. This "variable order" metastasis progression model would be in contrast to current models of cancer progression, in which events in tumorigenesis are proposed to occur in the same sequence every time a tumor arises.

Further studies, especially using multiple primary tumors and metastases from patients with different clinical histories, will be needed to clarify the role of specific genes and their cellular pathways to the overall metastatic potential. Our results on lymph node metastases should also be extended to distant metastases to identify genes responsible for the establishment and maintenance of metastasis.

In conclusion, our data provide experimental evidence for the model recently proposed by Hynes[53] that postulates the development of metastatic variants from a metastasis-prone primary tumor cell population. Moreover, our data extend this model by indicating that the sum of the changes in pathways and processes, rather than the specific genes or the order in which changes occur, determines the final metastatic outcome.

REFERENCES

1. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100:57–70.
2. Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. Nat Rev Cancer 2002;2:563–572.
3. Gupta PB, Mani S, Yang J, Hartwell K, Weinberg RA. The evolving portrait of cancer metastasis. Cold Spring Harb Symp Quant Biol 2005;70:291–297.
4. Hahn WC, Weinberg RA. Rules for making human tumor cells. N Engl J Med 2002;347:1593–1603.
5. Fidler IJ. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. Nat Rev Cancer 2003;3:453–458.
6. Fidler IJ, Kripke ML. Metastasis results from preexisting variant cells within a malignant tumor. Science 1977;197:893–895.
7. Bernards R, Weinberg RA. A progression puzzle. Nature 2002;418:823.
8. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530–536.
9. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002;347:1999–2009.
10. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. Nat Genet 2003;33:49–54.
11. Buckhaults P. Gene expression determinants of clinical outcome. Curr Opin Oncol 2006;18:57–61.
12. Califano J, van der Riet P, Westra W, et al. Genetic progression model for head and neck cancer: implications for field cancerization. Cancer Res 1996;56:2488–2492.
13. Forastiere A, Koch W, Trotti A, Sidransky D. Head and neck cancer. N Engl J Med 2001;345:1890–1900.
14. Perez-Ordonez B, Beauchemin M, Jordan RC. Molecular biology of squamous cell carcinoma of the head and neck. J Clin Pathol 2006;59:445–453.
15. Mao L, Hong WK, Papadimitrakopoulou VA. Focus on head and neck cancer. Cancer Cells 2004;5:311–316.
16. Lothaire P, de Azambuja E, Dequanter D, et al. Molecular markers of head and neck squamous cell carcinoma: promising signs in need of prospective evaluation. Head Neck 2006;28:256–269.
17. Ferlito A, Rinaldo A, Devaney KO, et al. Prognostic significance of microscopic and macroscopic extracapsular spread from metastatic tumor in the cervical lymph nodes. Oral Oncol 2002;38:747–751.
18. Li C, Wong HW. Model-based analysis of oligonucleotide arrays: model validation, design issues and standard error application. Genome Biol 2001;2:1–11.
19. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proc Natl Acad Sci U S A 2001;98:31–36.
20. Lemon WJ, Palatini JJ, Krahe R, Wright FA. Theoretical and experimental comparisons of gene expression indexes for oligonucleotide arrays. Bioinformatics 2002;18:1470–1476.
21. Morris JS, Yin G, Baggerly KA, Wu C, Zhang L. Pooling information across different studies and oligonucleotide chip types to identify prognostic genes for lung cancer. In: Shoemaker JS, Lin JS, editors. Methods of microarray data analysis IV. New York: Springer-Verlag; 2005. pp 51–66.
22. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249–264.
23. Dong Y, Sui L, Watanabe Y, Sugimoto K, Tokuda M. Clinical relevance of cyclin B1 overexpression in laryngeal squamous cell carcinoma. Cancer Lett 2002;177:13–19.
24. Hassan KA, Ang KK, El-Naggar AK, et al. Cyclin B1 overexpression and resistance to radiotherapy in head and neck squamous cell carcinoma. Cancer Res 2002;62:6414–6417.
25. Xi S, Dyer KF, Kimak M, et al. Decreased STAT1 expression by promoter methylation in squamous cell carcinogenesis. J Natl Cancer Inst 2006;98:181–189.
26. Pramana J, Pimentel N, Hofland I, et al. Heterogeneity of gene expression profiles in head and neck cancer. Head Neck 2007:1083–1089.
27. Lu Y, Lemon W, Liu PY, et al. A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. PLoS Medicine 2006;3:2229–2243.
28. Ramesh R, Ito I, Gopalan B, Saito Y, Mhashilkar AM, Chada S. Ectopic production of MDA-7/IL-24 inhibits invasion and migration of human lung cancer cells. Mol Ther 2004;9:510–518.
29. Zlotnik A. Chemokines and cancer. Int J Cancer 2006;119:2026–2029.

30. Levine MD, Liotta LA, Stracke ML. Stimulation and regulation of tumor cell motility in invasion and metastasis. EXS 1995;74:157–179.
31. Chung CH, Parker JS, Karaca G, et al. Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. Cancer Cells 2004;5:489–500.
32. Cromer A, Carles A, Millon R, et al. Identification of genes associated with tumorigenesis and metastatic potential of hypopharyngeal cancer by microarray analysis. Oncogene 2004;23:2484–2498.
33. Roepman P, Wessels LF, Kettelarij N, et al. An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. Nat Genet 2005;37:182–186.
34. O'Donnell RK, Kupferman M, Wei SJ, et al. Gene expression signature predicts lymphatic metastasis in squamous cell carcinoma of the oral cavity. Oncogene 2005;24:1244–1251.
35. Schmalbach CE, Chepeha DB, Giordano TJ, et al. Molecular profiling and the identification of genes associated with metastatic oral cavity/pharynx squamous cell carcinoma. Arch Otolaryngol Head Neck Surg 2004;130:295–302.
36. Hunter KD, Thurlow JK, Fleming J, et al. Divergent routes to oral cancer. Cancer Res 2006;66:7405–7413.
37. Roepman P, de Jager A, Groot Koerkamp MJ, Kummer JA, Slootweg PJ, Holstege FC. Maintenance of head and neck tumor gene expression profiles upon lymph node metastasis. Cancer Res 2006;66:11110–11114.
38. Giri U, Ashorn CL, Ramdas L, et al. Molecular signatures associated with clinical outcome in patients with high-risk head-and-neck squamous cell carcinoma treated by surgery and radiation. Int J Radiat Oncol Biol Phys 2006;64:670–677.
39. Sotiriou C, Lothaire P, Dequanter D, Cardoso F, Awada A. Molecular profiling of head and neck tumors. Curr Opin Oncol 2004;16:211–214.
40. Weigelt B, Glas AM, Wessels LF, Witteveen AT, Peterse JL, van't Veer LJ. Gene expression profiles of primary breast tumors maintained in distant metastases. Proc Natl Acad Sci U S A 2003;100:15901–15905.
41. Belbin TJ, Singh B, Smith RV, et al. Molecular profiling of tumor progression in head and neck cancer. Arch Otolaryngol Head Neck Surg 2005;131:10–18.
42. Forastiere A, Koch W, Trotti A, Sidransky D. Head and neck cancer. N Engl J Med 2001;345:1890–1900.
43. Nam SW, Clair T, Campo CK, Lee HY, Liotta LA, Stracke ML. Autotaxin (ATX), a potent tumor motogen, augments invasive and metastatic potential of ras-transformed cells. Oncogene 2000;19:241–247.
44. Kang Y, Siegel PM, Shu W, et al. A multigenic program mediating breast cancer metastasis to bone. Cancer Cell 2003;3:537–549.
45. Tsuda H, Sekine K, Fujita K, Ligo M. Cancer prevention by bovine lactoferrin and underlying mechanisms—a review of experimental and clinical studies. Biochem Cell Biol 2002;80:131–136.
46. Kyriazanos ID, Tachibana M, Dhar DK, et al. Expression and prognostic significance of S100A2 protein in squamous cell carcinoma of the esophagus. Oncol Rep 2002;9:503–510.
47. Ellerhorst JA, Prieto VG, Ekmekcioglu S, et al. Loss of MDA-7 expression with progression of melanoma. J Clin Oncol 2002;20:1069–1074.
48. Rubin H. Selected cell and selective microenvironment in neoplastic development. Cancer Res 2001;61:799–807.
49. Matrisian LM, Cunha GR, Mohla S. Epithelial-stromal interactions and tumor progression: meeting summary and future directions. Cancer Res 2001;61:3844–3846.
50. Liotta LA, Kohn EC. The microenvironment of the tumour–host interface. Nature 2001;411:375–379.
51. Clark EA, Golub TR, Lander ES, Hynes RO. Genomic analysis of metastasis reveals an essential role for RhoC. Nature 2000;406:532–535.
52. Chen Z, Zhang K, Zhang X, et al. Comparison of gene expression between metastatic derivatives and their poorly metastatic parental cells implicates crucial tumor-environment interaction in metastasis of head and neck squamous cell carcinoma. Clin Exp Metastasis 2003; 20:335–342.
53. Hynes RO. Metastatic potential: generic predisposition of the primary tumor or rare, metastatic variants-or both? Cell 2003;113:821–823.

Beyond Receptor Expression Levels: The Relevance of Target Accessibility in Ligand-Directed Pharmacodelivery Systems

Michael G. Ozawa, Amado J. Zurita, Emmanuel Dias-Neto,
Diana N. Nunes, Richard L. Sidman, Juri G. Gelovani,
Wadih Arap*,[1], and Renata Pasqualini*,[1]

For development of a new ligand-directed pharmacology, it is critical to measure delivery of targeted drug ligands via molecular imaging or diagnostic readouts (termed theranostics). Combinatorial peptide libraries serve as unbiased functional screens that can identify specific peptides targeting cell-surface receptors accessible to the circulation. As candidate drug leads, such peptides provide motifs likely to modify ligand-receptor interactions and downstream signal transduction pathways. This strategy is synergistic with genomic and proteomic approaches and has yielded insights into the specialized nature of the target tissue microenvironment. However, for this vision to be realized, one must look, as recent literature suggests, beyond receptor levels and critically analyze ligand accessibility as a key determinant in pharmacodelivery systems. (Trends Cardiovasc Med 2008;18:126–133) © 2008, Elsevier Inc.

Michael G. Ozawa, Amado J. Zurita, Emmanuel Dias-Neto, Diana N. Nunes, Wadih Arap, and Renata Pasqualini are at the Department of Genitourinary Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA; Department of Cancer Biology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA. Juri G. Gelovani is at the Department of Experimental Diagnostic Imaging, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA. Richard L. Sidman is at the Harvard Medical School and Department of Neurology, Beth Israel Deaconess Medical Center, Harvard Institutes of Medicine, Boston, MA 02115, USA.

* Address correspondence to: Renata Pasqualini or Wadih Arap, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, TX 77030, USA. Tel.: (+1) 713 792 3873; fax: (+1) 713 745 2999; e-mails: rpasqual@mdanderson.org, warap@mdanderson.org

[1] These authors contributed equally to this work.

## • Introduction and Challenge of Basic Premises

The much-heralded era of personalized medicine has the potential to bring together therapeutic and diagnostic strategies. Under this vision, each drug will have a companion diagnostic blood or imaging test to help treating physicians predict therapeutic response and/or monitor disease progression; in principle, such merging (often referred to as theranostics) may encompass specific serum or tissue biomarkers, pharmacogenomics, and molecular-genetic imaging studies, either as single agents or in combination (Warner 2004). Despite concerns and uncertainty over definitions, terminology, and cost, the US Food and Drug Administration has recently put forward the concept of "drug-diagnostic co-development" with guidelines to promote effective progress in this area (Hinman et al. 2006). In oncology, possible theranostic scenarios might comprise, for example, measuring tissue expression of receptors such as Her-2/Neu for monoclonal antibodies such as herceptin (Pritchard et al. 2006), predicting the aggressive behavior of tumors by use of RNA microarrays (Van 't Veer et al. 2002), or detecting epithelial growth factor receptor mutations that determine tumor resistance to small molecules such as gefitinib (Kobayashi et al. 2005).

Successful development of ligand-directed pharmacology in an era of theranostics will require thorough characterization of targeting ligands and their corresponding receptors, as well as a deep understanding of the resultant intracellular signaling cascades that yield a physiologic or pathologic readout. Although a large number of methods are available to identify novel ligands, receptors, or both, prioritization of candidates remains a significant hurdle, for poor leads can create an enormous financial and time expenditure. New tactics are surfacing to streamline the discovery process. Recently, Cheng et al. described a new computational method to assess potential targetability or druggability by affinity analysis of receptor structure (Cheng et al. 2007). This approach helps predict molecular targets suitable for therapeutic exploitation and potentially reduces pursuit of nonoptimal receptors, but it also highlights the pressing need for screening and validation approaches at multiple steps along the developmental pipeline and for integrative approaches for identification and candidate pursuit.

Furthermore, it should logically follow that—beyond protein expression levels in tissues and mutational and/or epigenetic events—a central (yet often unrecognized or overlooked) aspect of directed therapies is the accessibility of target receptors to circulating ligands. Accessibility assumes particular importance if the target is a cell-membrane-associated receptor whose expression is restricted to the endothelium of certain organs or is up-regulated in diseases featuring abnormal angiogenesis (such as cancer, proliferative retinopathies,

Figure 1. Ligand-directed vascular targeting in normal and tumor settings. **(A)** Endothelial cells lining blood vessels in normal tissues form a tight physiologic barrier for circulating ligands and enable little or no passive extravasation. **(B)** In contrast, the vasculature in tumors is marked by an abnormal endothelium that enables "leakage" and binding of circulating ligands to extravascular receptor targets.

rheumatoid arthritis, and others). A basic tenet of vascular targeting, by circulating ligands, is that the first cell layer a systemically administered agent will encounter is the vascular endothelium (Figure 1); the highly desired access to "outer" cell layers may also occur but depends on other variables such as circulation time, tissue-specific endothelial permeability, differential hydrostatic pressures, ligand/receptor-mediated internalization by endothelial cells, on/off target receptor rates, and activity of protein degradation mechanisms (Carmeliet 2003, Hajitou et al. 2006a, Neri and Bicknell 2005). These parameters (and perhaps others still unknown) will determine the net inside-out accessibility to endothelial, stromal, and parenchymal cells.

By screening combinatorial peptide libraries on cells in vitro and in intact experimental animals and patients, our group—among others—has identified several tissue-specific and angiogenesis-related vascular "ZIP codes," cell-surface receptors that enable ligand-directed homing through the circulation. Among these cell-surface receptors are certain integrins (Arap et al. 1998, Hajitou et al. 2006b), matrix metalloproteases (Koivunen et al. 1999), vascular endothelial growth factor receptors (Giordano et al. 2005), cytokine receptors (Su et al.

2005, Zurita et al. 2004), stress-response chaperones (Arap et al. 2004), membrane-bound aminopeptidases (Essler and Ruoslahti 2002, Marchiò et al. 2004, Pasqualini et al. 2000), and even proteoglycans (Burg et al. 1999). Here we discuss the versatility of phage display-based technology as a functional ligand discovery platform, the diversity of the corresponding receptors, and concepts for future development.

- **Selection of Ligand–Receptors in a Functional Context**

From a purely genomics perspective, the identification of nucleic acid sequences encoding genes over-expressed in cancer (or in other human diseases) is typically the first step in the development of targeting strategies. To find such relevant molecules, paired cell or tissue samples of normal and tumor origin are analyzed. High-throughput screening methodologies include complementary DNA (cDNA) microarrays, generation of expressed sequence tags (ESTs), serial analysis of gene expression (SAGE), and massive parallel sequence signatures (MPSS) (Adams et al. 1992, Brenner et al. 2000, DeRisi et al. 1996, St Croix et al. 2000). Whereas cDNA microarrays enable the quantification of messenger RNAs through the use of cDNA probes

over a predefined population of spotted transcripts, the remaining techniques rely on cDNA sequencing to determine the frequency of the active transcripts in the biologic sample of interest. With this approach, differential transcript expression was identified in human cancer; moreover, although technically challenging, such analysis can be extended to subpopulations within the malignant compartment (i.e., tumor cells and possibly cancer stem cells) and nonmalignant compartment (i.e., stromal and endothelial cells) (St Croix et al. 2000).

Although genomic studies and related proteomic studies do narrow the large list of molecular signatures of disease, identification of viable candidates can remain difficult owing to large sets of potential receptors, complex cellular compositions of the tissue samples, translational and posttranslational modifications, and changes in protein localization in gene expression studies. It is often unsolved whether identified targets are uniquely localized to the cell surface in disease or even whether they contribute sufficiently to enable selective systemic targeting. Alternative methods for organ- or disease-specific targeting have been developed to help circumvent these uncertainties, including recent work that used

circulating activated biotin as a screening tool for receptor identification (Roesli et al. 2006). However, although the above-mentioned approaches yield potential molecular receptors or targets, they do not directly establish specific and clinically effective ligand/receptor pairings.

Thus, profiling techniques that address cell-surface localization and accessibility while simultaneously identifying targeting sequences would be potentially advantageous. In essence, there are two main approaches for uncovering vascular targeting ligand–receptor systems based on direct protein interactions: in vitro or in vivo selection of combinatorial libraries and generation of monoclonal antibodies against antigens expressed in the endothelium; the latter was recently exploited for the development of novel tumor targeting moieties by fusion of monoclonal antibody derivates (Wu et al. 2007). However, although there are critical elements and main conclusions that are shared by this and other targeting methodologies, we focus predominantly on the use of phage display as a functional screening method for the development of ligand-directed targeting systems.

Phage display technology was originally designed to identify binding sites of antibodies ("epitope mapping") (Scott and Smith 1990). This technique has since been greatly expanded to enable screening of the antibody repertoire in serum, the receptor diversity in isolated cell populations, and the heterogeneity of the vascular endothelium in animal models and in patients (Arap et al. 2002, Hajitou et al. 2006a, Pentz et al. 2005). In cancer, phage clones binding to membrane receptors often carry peptides that target readily accessible receptors, be they on cell surfaces of the tumor vascular endothelium, associated stromal elements, malignant cells per se, on some combination of these cellular compartments, and even on extracellular matrix. For the in vivo system, phage libraries display 10^9 or more unique peptide sequences on the minor coat protein to probe the receptor repertoire of cells in various organs of interest (Arap et al. 2004, Arap et al. 1998). Enrichment of selective peptides can be monitored by DNA sequence analysis through multiple rounds of selection by comparing the frequency

of repeated peptide sequences to both unselected library and control tissues. Analysis of peptides is carried out with basic local alignment search tool and ClustalW, along with other "off-the-shelf" and/or "customized" data mining applications that can identify proteins with which the targeting peptides share sequence similarity and can align them to putative sites. This type of approach was used for several vascular receptors and their corresponding targeted peptide ligands (Arap et al. 2002, Koivunen et al. 1999, Marchiò et al. 2004, Pasqualini et al. 2000). Alternatively (or in addition to online similarity searches), classic bioche!mistry methodology such as standard affinity chromatography also can be used (Giordano et al. 2005, Kolonin et al. 2004, Rajotte and Ruoslahti 1999).

• **Cell-Bound Aminopeptidases as Prototypic Vascular Targets**

Remarkably, although the requirement of receptor accessibility in ligand-directed pharmacodelivery would appear obvious (or at least intuitive), accessibility has been relatively little explored in a systematic manner. Critics and commentators have pointed out that the broad expression patterns of the receptors found by phage display technology might often render them unsuitable as targets for ligand-directed intervention. As an illustrative proof-of-concept to address this scientific challenge and to advance knowledge about ligand-directed targeting, we focused on certain cell-membrane-associated aminopeptidases for several reasons (Marchiò et al. 2004, Pasqualini et al. 2000). First, we have recently proposed disciplined and specific criteria to define ZIP codes accessible to the vascular endothelial cells, perivascular cells (pericytes), and even tumor cells in structural and mechanistic terms. Second, among the handful of targeted ligand–receptor systems uncovered by screening peptide libraries (Arap et al. 1998, Koivunen et al. 1999, Zurita et al. 2004) and also by other vascular targeting strategies such as the hybridoma-free generation of monoclonal antibodies against tumor or tumor-associated antigens, at least three different aminopeptidases (aminopeptidase A [APA], aminopeptidase N

[APN], and aminopeptidase P [APP]) have shown proven efficacy for pharmacodelivery in preclinical settings (Essler and Ruoslahti 2002, Marchiò et al. 2004, Pasqualini et al. 2000). Third, nonmutually exclusive alternative mechanistic explanations involving conformational active states ("phenotypic activation") in tumors vs normal tissues have also been invoked in this matter (Curnis et al. 2002). Fourth, genetic knockouts have been made; and APA-null (Mitsui et al. 2003) and APN-null (Rangel et al. 2007) mice have been evaluated as tumor vascular ZIP codes. Fifth, although the cell-membrane-bound aminopeptidases are generally considered "ubiquitous" and often deemed poor ligand-directed delivery systems, they continually are identified as vascular targets in functional screenings outlined here (Essler and Ruoslahti 2002, Marchiò et al. 2004, Pasqualini et al. 2000). Therefore, how can one resolve this apparent paradox of receptor expression vs target accessibility?

The aminopeptidases are a large family of enzymes involved in the maturation, regulation, and degradation of proteins. Aminopeptidase N (CD13, E.C. 3.4.11.2; encoded by the *ANPEP* gene) is a zinc metalloproteinase that specifically cleaves N-terminal amino acid residues with neutral side chains (Look et al. 1989). Angiotensin III is the only known native substrate of APN, cleaving angiotensin III into angiotensin IV, which participates in the renin–angiotensin cascade by participating in neuronal signaling, as well as in the degradation of peptides in the intestinal brush border (Sjostrom et al. 2000). In contrast, APA (E.C. 3.4.11.7; encoded by the *ENPEP* gene) cleaves N-terminal glutamyl or aspartyl residues from polypeptides. Like APN, APA also participates in the renin–angiotensin system by converting angiotensin II to angiotensin III and also has links to the immune system via involvement in B-cell differentiation (Li et al. 1993). As a third example, APP (E.C. 3.4.11.9; encoded by the *XPNPEP2* gene) is a GPI-linked membrane protein expressed on the surface of vascular endothelial and lymphoid cells of various tissues (Lasch et al. 1998). This widely distributed hydrolase is specific for N-terminal imido bonds, which are common to several collagen degradation products,

neuropeptides, vasoactive peptides, and cytokines (Lendeckel et al. 2000).

Phage display-based technology played a central role in establishing aminopeptidases as relevant molecular targets in angiogenesis (Marchiò et al. 2004, Pasqualini et al. 2000, Rangel et al. 2007), and we have indeed more broadly hypothesized that the renin–angiotensin system (hence APA and APN) may have a regulatory role in cancer-associated angiogenesis (Khakoo A et al. unpublished data). A functional role for APN in angiogenesis and tumor growth was noted when the tripeptide motif asparagine–glycine–arginine (NGR, single-letter code), recovered among the dominant peptide sequences during a combinatorial peptide library test in tumor xenograft models, was shown to localize strongly and specifically to tumor blood vessels (Pasqualini et al. 2000). The corresponding receptor for the NGR motif was APN, which is present in angiogenic blood vessels (Pasqualini et al. 2000). We also showed that coupling the NGR motif to a cytotoxin (Arap et al. 1998) or fusing it to a synthetic proapoptotic peptidomimetic (Ellerby et al. 1999) increased antitumor efficacy compared with the nontargeted parental compounds; other groups observed similar effects by incorporating multiple NGR motifs into different recombinant cytokines (Curnis et al. 2002, Curnis et al. 2005) in preclinical models and quite recently in clinical trial settings (Bordignon C et al. unpublished). More recently, we demonstrated in APN-null mice that this enzyme affects pathologic angiogenesis while sparing embryonic vasculogenesis and physiologic angiogenesis (Rangel et al. 2007). Similarly, we showed that APA-binding peptides—such as the motif CPRECES—selectively targeted perivascular cells (pericytes) in angiogenic blood vessels and had potent antiangiogenic properties (Marchiò et al. 2004). Thus, we rationally designed studies to target this vascular receptor because APA is upregulated and functional in activated pericytes of tumor blood vessels. With an entirely different phage display system in vivo, other investigators identified APP as a receptor selectively present in vasculature of normal and cancerous breast tissue in mice (Essler and Ruoslahti 2002).

Given our working hypothesis that phage display-based combinatorial selections in vivo may yield complementary, confirmatory, or at times a different set of targets than genomic or proteomic strategies, we asked what factors might form the basis for selectivity of receptors for targeted ligands. Among the possibilities are primary systemic accessibility via the circulation, in vivo microanatomic context, receptor-mediated internalization, or even known or as yet undetermined environment-dependent receptor features. Thus, one might conjecture that the membrane-bound aminopeptidases might have been recognized as potential tumor vascular targets by use of genomics-based receptor discovery platforms (rather than selection by in vivo phage display).

To that end, we evaluated the total set of human SAGE tags (18.7 million) and ESTs (5.9 million) available in the Cancer Genome Anatomy Project database (http://cgap.nci.nih.gov) and also performed MPSS analysis (http://mpss.licr.org) for gene expression of the aminopeptidases described above. Based on such analyses, we noted a relatively higher expression of APN (487 SAGE tags, 383 ESTs), lower expression of APA (27 SAGE tags, 148 ESTs), and a rare expression of APP (four SAGE tags, 29 ESTs); a similar pattern of gene expression also emerged from MPSS.

For a more detailed evaluation of gene expression of these three enzymes in tumor and nontumor tissues, we focused our analysis on SAGE tags produced for selected tissues ($n = 8$) for which tumor and nontumor sequences are available (Figure 2). From SAGE data, APA is abundantly expressed in the vasculature and kidney, whereas APN showed high expression in kidney, colon, breast, prostate, and vasculature. With respect to tumor vs nontumor tissues, APA appears down-regulated in kidney and liver tumors and up-regulated in lung and brain tumors, as well as in tumor vasculature. By contrast, the data suggest that APN is up-regulated in brain and lung tumors but downregulated in cancers of colon, kidney, and liver, along with tumor vasculature. No appreciable expression variation was observed for APA or APN between nontumor/tumor pairs in prostate or breast.

We have previously established that endothelial cells and perivascular cells (presumably pericytes) in human breast carcinoma express APN, whereas blood vessels in normal breast tissue are essentially negative (Pasqualini et al. 2000). In contrast, SAGE data illustrates positive expression of APN and no significant difference in transcripts between normal and tumor tissue (Figure 2). These data are further corroborated by LongSAGE data, where analysis of 9×10^5 tags from breast tissue showed an average of 5.02 tags per 2×10^5 tags in nontumor tissue and 5.3 per 2×10^5 tags in cancer samples (data not shown). LongSAGE provides greater confidence in consistent identification because probes are larger and more specific (Saha et al. 2002). Notably, these SAGE libraries came from bulk tissue; and a very different picture might have emerged if the vasculature had been microdissected so that its specific markers would not have been "diluted" among genes of the tissues as a whole. Finally, APP could only be detected in the most sequenced tissues and with too few tags through this approach to enable meaningful conclusions. This analysis clearly indicates that although the transcriptional program may largely define cells and tissues in terms of morphology and physiology, the identification of markers based on transcriptional comparison of all genes in a biologic sample is markedly complicated by the extreme abundance of a few transcripts, leaving the less abundant genes far below the significance cutoff.

In conclusion, although there are relevant factors to consider in SAGE tag analysis, such as tissue- and cell-type contributions, sample numbers, SAGE library availability, and total tags generated, the aggregate of the data point to ubiquitous gene expression patterns of aminopeptidases and only modest expression changes between nontumor/tumor pairs. Thus, it is unlikely that one would identify membrane-bound aminopeptidases as candidate targets by these genetic approaches in isolation. Such conclusions are in line with the reported widespread expression of these proteins, hence the apparent paradox (Li et al. 1993, Look et al. 1989).
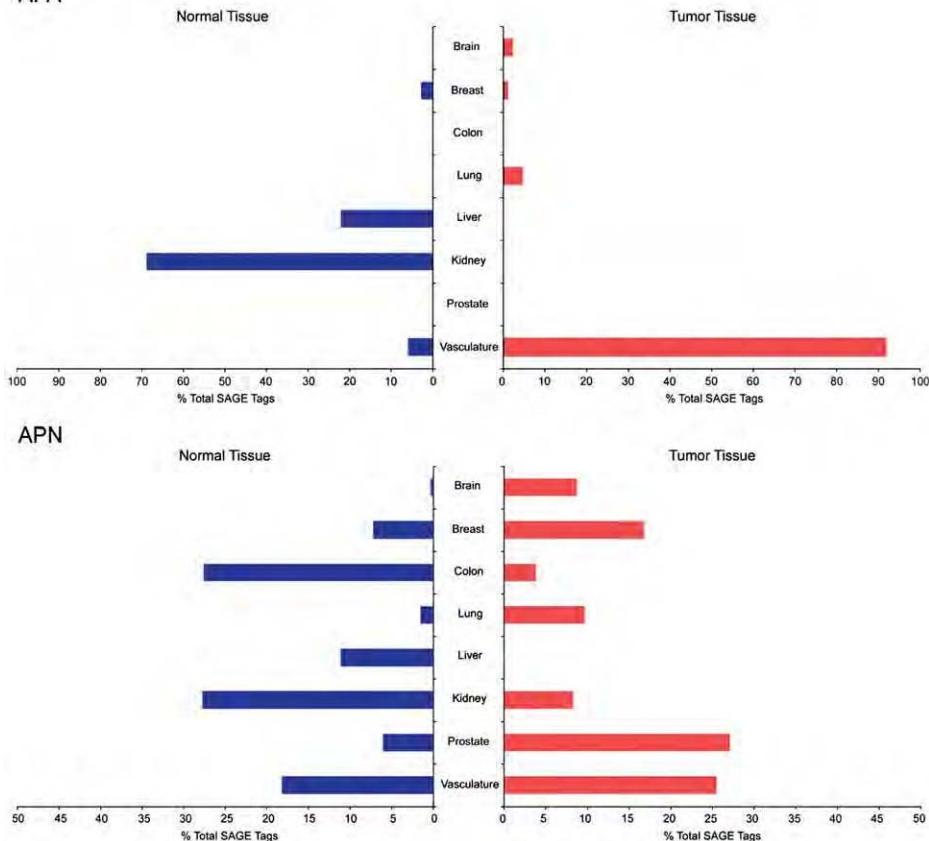
Figure 2. Genomic analysis of aminopeptidase expression. APA and APN transcripts are used. The SAGE data obtained from nontumor and tumor samples are depicted. Values are normalized as percentage of tags identified from the total pool of analyzed SAGE tags.

- **Receptor Expression Level vs Accessibility to Circulating Ligands: Critically Reconciling a Vascular Targeting Paradox**

Reconciliation of the data on gene expression levels with the data obtained by in vivo phage display in preclinical models is possible through a simple validation assay after the distribution of intravenously administered specific antibodies. Endothelial cells lining the vascular system are known to influence protein expression within the tissue microenvironment (Lammert et al. 2001). These cells form the primary physiologic barrier between blood and the tissue surrounding blood vessels. Thus, in a reductionist view, agents circulating in the bloodstream might readily target luminally expressed surface receptors on normal endothelial cells and would not have immediate access to cell-surface receptors expressed in deeper tissue layers. In contrast, at sites in which the vascular system is abnormally structured or has lost its normal physiologic barrier function—such as in the angiogenic vasculature in

cancer—targets outside of blood vessels might be accessible and successfully targeted (Carmeliet 2003, Marchiò et al. 2004, Pasqualini et al. 2000).

Supporting the SAGE tag analysis, immunohistochemical detection of cell-bound APN (Figure 3) indicates strong immunoreactivity in the renal tubules of the normal mouse kidney (Figure 3A) and in the pericytes lining brain capillaries (Figure 3B). However, after intravenous administration of the same antibody against APN and a 24-hour rather than the more conventional ~5-min circulation, immunohistochemical analysis did not detect presence of the antibody within the kidney (Figure 3C) or the brain (Figure 3D). Consistent with a previous report (Mentzel et al. 1999), we have obtained similar results with anti-APA antibodies.

In previous work, we have shown that both APA and APN expressions localize to pericytes lining blood vessels in tumors (Marchiò et al. 2004, Pasqualini et al. 2000), confirmed through confocal microscopy and colocalization with the established pericyte markers NG2 proteoglycan, desmin, and αSMA (Figure 2E-G) (Morikawa et al. 2002). After systemic (intrave-

nous) administration of specific antibodies against APA or APN, the distribution of detectable APA and APN parallels what is seen with standard immunohistochemical detection (Figure 2H and I). Whereas kidney and brain express high levels of both targets, the vasculature shields parenchymal cells that may express the target from being reached by the probes. In contrast, the "leaky" vasculature within tumors enables extravasation and binding to target-expressing cells, including parenchymal and immune cells (Carmeliet 2003, Neri and Bicknell 2005). These results point to a clear advantage of functional screening assays for target identification and validation, relative to current "omics" methodologies. Vascular endothelial cells and pericytes account for a very minor percentage of the total cell population, as evidenced by the small contribution of APA and APN transcripts to the overall transcript levels (Figure 2). Yet, screening with in vivo phage display did identify these novel regulators of angiogenesis. Together, these data support the significance of target accessibility as a central factor in our understanding of normal and tumor cell behavior.
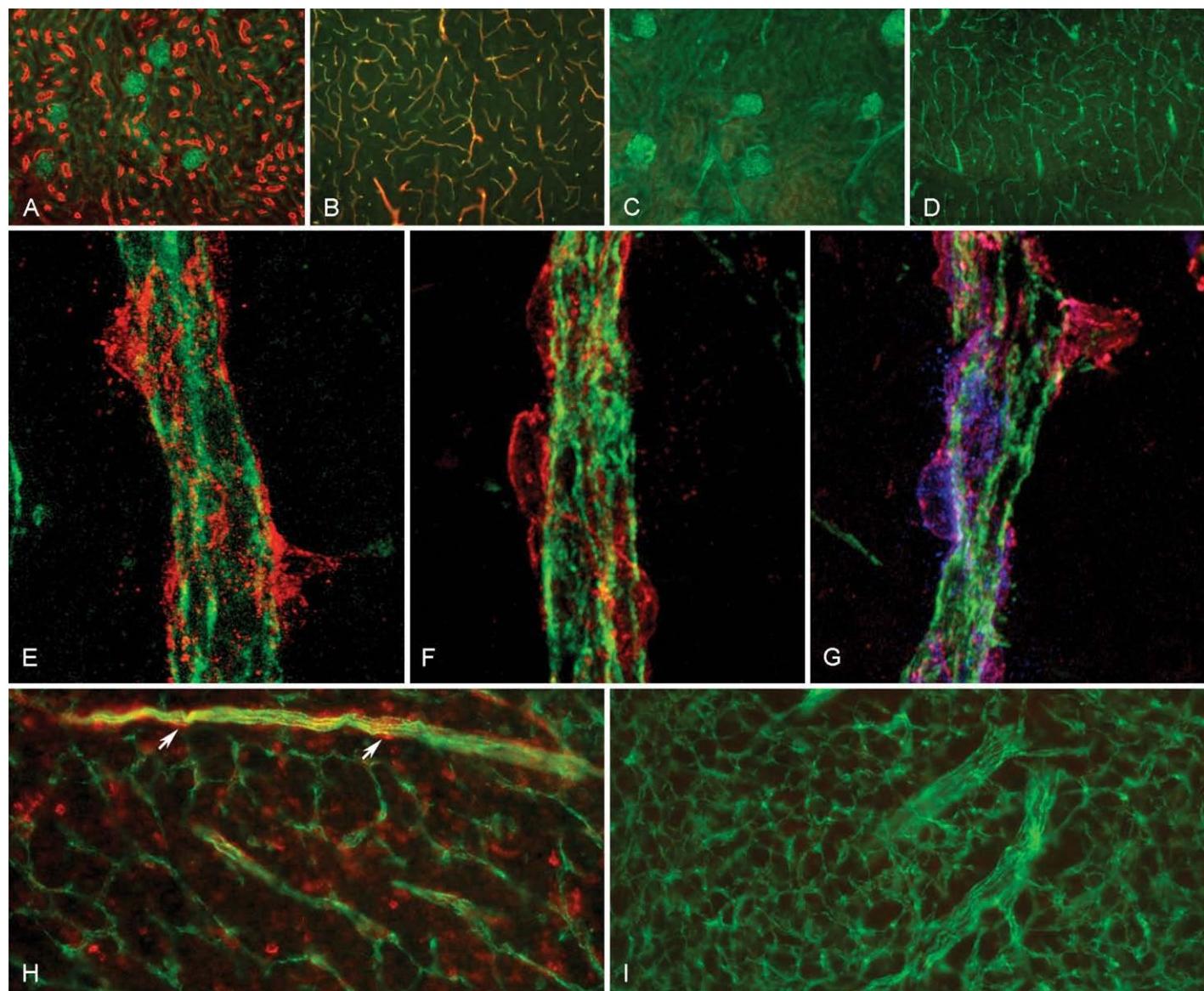
Figure 3. Protein expression vs accessibility of aminopeptidases. **(A, B)** Double-immunofluorescence images identify blood vessels (*green*) and APN immunoreactivity (*red*) in normal kidney tubules and in pericytes lining normal brain capillaries. **(C, D)** After systemic (intravenous) administration of specific anti-APN antibodies (clone R3-63; AbD Serotec), APN immunoreactivity (*red*) is not detected in the kidney or brain. **(E-G)** Confocal micrographs in an isogenic mouse mammary tumor (Hajitou et al. 2006b) established the localization of APA expression **(E**, *red*) to pericytes identified by expression of the NG2 pericyte marker **(F**, *red*) associated with tumor blood vessels (*green*); APA (*blue*) and NG2 (*red*) colocalize in pericytes **(G)**. **(H, I)** Intravenous administration of anti-APN antibodies illustrates accessibility of APN expressed on extravascular cells **(H**, *red*) including pericytes **(H**, *red*, *arrow*) associated with tumor blood vessels (*green*), whereas a control immunoglobulin G is minimally detected **(I**, *red*).

• Conclusions and Future Directions

Future development of new pharmacologic agents based on peptide ligands requires the integration of therapeutics and diagnostics into theranostics. Reliable and sensitive methodologies that enable targeted molecular imaging are essential for rational targeted drug development. To that end, the selection of combinatorial phage display libraries in vivo can be used as functional screens to identify selective and accessible vascular ligand/receptor systems. The use of novel ligand-directed hybrid vectors such as adeno-associated virus phage (Hajitou et al. 2006b) and combinations of targeting peptidomimetics and cytotoxic agents (Arap et al. 1998) or radiologically visualizable agents will enable relatively selective destruction of pathologic vasculature and tumors, and molecular-genetic imaging of angiogenesis-related or tissue-specific vascular endothelium markers. Moreover, these strategies may dovetail to enable compressed drug development timelines and greater efficiency through "phase zero" trials conducted under the Exploratory Investigational New Drug Guidance from the Food and Drug Administration (Kummar et al. 2007).

References

Adams MD, Dubnick M, Kerlavage AR, et al: 1992. Sequence identification of 2,375 human brain genes. Nature 355: 632–634.

Arap MA, Lahdenranta J, Mintz PJ, et al: 2004. Cell surface expression of the stress response chaperone GRP78 enables tumor targeting by circulating ligands. Cancer Cell 6:275–284.

Arap W, Kolonin MG, Trepel M, et al: 2002. Steps toward mapping the human vasculature by phage display. Nat Med 8:121–127.

Arap W, Pasqualini R, Ruoslahti E: 1998. Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. Science 279:377–380.

Brenner S, Johnson M, Bridgham J, et al: 2000. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. Nat Biotechnol 18: 630–634.

Burg MA, Pasqualini R, Arap W, Ruoslahti E Stallcup WB: 1999. NG2 proteoglycan-binding peptides target tumor neovasculature. Cancer Res 59:2869–2874.

Carmeliet P: 2003. Angiogenesis in health and disease. Nat Med 9:653–660.

Cheng AC, Coleman RG, Smyth KT, et al: 2007. Structure-based maximal affinity model predicts small-molecule druggability. Nat Biotechnol 25:71–75.

Curnis F, Arrigoni G, Sacchi A, et al: 2002. Differential binding of drugs containing the NGR motif to CD13 isoforms in tumor vessels, epithelia, and myeloid cells. Cancer Res 62:867–874.

Curnis F, Gasparri A, Sacchi A, et al: 2005. Targeted delivery of IFN gamma to tumor vessels uncouples antitumor from counter-regulatory mechanisms. Cancer Res 65: 2906–2913.

DeRisi J, Penland L, Brown PO, et al: 1996. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat Genet 14:457–460.

Ellerby HM, Arap W, Ellerby LM, et al: 1999. Anti-cancer activity of targeted pro-apoptotic peptides. Nat Med 5:1032–1038.

Essler M, Ruoslahti E: 2002. Molecular specialization of breast vasculature: a breast-homing phage-displayed peptide binds to aminopeptidase P in breast vasculature. Proc Natl Acad Sci U S A 99: 2252–2257.

Giordano RJ, Anobom CD, Cardó-Vila M, et al: 2005. Structural basis for the interaction of a vascular endothelial growth factor mimic peptide motif and its corresponding receptors. Chem Biol 12:1075–1083.

Hajitou A, Pasqualini R, Arap W: 2006a. Vascular targeting: recent advances and therapeutic perspectives. Trends Cardiovasc Med 16:80–88.

Hajitou A, Trepel M, Lilley CE, et al: 2006b. A hybrid vector for ligand-directed tumor targeting and molecular imaging. Cell 125: 385–398.

Hinman LM, Huang SM, Hackett J, et al: 2006. The drug diagnostic co-development concept paper. Commentary from the 3rd FDA-DIA-PWG-PhRMA-BIO Pharmacogenomics Workshop. Pharmacogenomics J 6:375–380.

Kobayashi S, Boggon TJ, Dayaram T, et al: 2005. EGFR mutation and resistance of non-small-cell lung cancer to gefitinib. N Engl J Med 352:786–792.

Koivunen E, Arap W, Valtanen H, et al: 1999. Tumor targeting with a selective gelatinase inhibitor. Nat Biotechnol 17:768–774.

Kolonin MG, Saha PK, Chan L, et al: 2004. Reversal of obesity by targeted ablation of adipose tissue. Nat Med 10:625–632.

Kummar S, Kinders R, Rubinstein L, et al: 2007. Compressing drug development timelines in oncology using phase '0' trials. Nat Rev Cancer 7:131–139.

Lammert E, Cleaver O, Melton D: 2001. Induction of pancreatic differentiation by signals from blood vessels. Science 294: 564–567.

Lasch J, Moschner S, Sann H, et al: 1998. Aminopeptidase P—a cell-surface antigen of endothelial and lymphoid cells: catalytic and immuno-histotopical evidences. Biol Chem 379:705–709.

Lendeckel U, Kahne T, Riemann D, et al: 2000. Review: the role of membrane peptidases in immune functions. Adv Exp Med Biol 477:1–24.

Li L, Wu Q, Wang J, et al: 1993. Widespread tissue distribution of aminopeptidase A, an evolutionarily conserved ectoenzyme recognized by the BP-1 antibody. Tissue Antigens 42:488–496.

Look AT, Ashmun RA, Shapiro LH, Peiper SC: 1989. Human myeloid plasma membrane glycoprotein CD13 (gp150) is identical to aminopeptidase N. J Clin Invest 83: 1299–1307.

Marchiò S, Lahdenranta J, Schlingemann RO, et al: 2004. Aminopeptidase A is a functional target in angiogenic blood vessels. Cancer Cell 5:151–162.

Mentzel S, Van Son JP, Dijkman HB, et al: 1999. Induction of albuminuria in mice: synergistic effect of two monoclonal antibodies directed to different domains of aminopeptidase A. Kidney Int 55: 1335–1347.

Mitsui T, Nomura S, Okada M, et al: 2003. Hypertension and angiotensin II hypersensitivity in aminopeptidase A-deficient mice. Mol Med 9:57–62.

Morikawa S, Baluk P, Kaidoh T, et al: 2002. Abnormalities in pericytes on blood vessels and endothelial sprouts in tumors. Am J Pathol 160:985–1000.

Neri D, Bicknell R: 2005. Tumour vascular targeting. Nat Rev Cancer 5:436–446.

Pasqualini R, Koivunen E, Kain R, et al: 2000. Aminopeptidase N is a receptor for tumor-homing peptides and a target for inhibiting angiogenesis. Cancer Res 60:722–727.

Pentz RD, Cohen CB, Wicclair M, et al: 2005. Ethics guidelines for research with the recently dead. Nat Med 11:1145–1149.

Pritchard KI, Shepherd LE, O'Malley FP, et al: 2006. HER2 and responsiveness of breast cancer to adjuvant chemotherapy. N Engl J Med 354:2103–2111.

Rajotte D, Ruoslahti E: 1999. Membrane dipeptidase is the receptor for a lung-targeting peptide identified by in vivo phage display. J Biol Chem 274: 11593–11598.

Rangel R, Sun Y, Guzman-Rojas L, et al: 2007. Impaired angiogenesis in aminopeptidase N-null mice. Proc Natl Acad Sci U S A 104:4588–4593.

Roesli C, Neri D, Rybak JN: 2006. In vivo protein biotinylation and sample preparation for the proteomic identification of organ- and disease-specific antigens accessible from the vasculature. Nat Protoc 1: 192–199.

Saha S, Sparks AB, Rago C, et al: 2002. Using the transcriptome to annotate the genome. Nat Biotechnol 20:508–512.

Scott JK Smith GP: 1990. Searching for peptide ligands with an epitope library. Science 249:386–390.

Sjostrom H, Noren O, Olsen J: 2000. Structure and function of aminopeptidase N. Adv Exp Med Biol 477:25–34.

St Croix B, Rago C, Velculescu V, et al: 2000. Genes expressed in human tumor endothelium. Science 289:1197–1202.

Su JL, Lai KP, Chen CA, et al: 2005. A novel peptide specifically binding to interleukin-6 receptor (gp80) inhibits angiogenesis and tumor growth. Cancer Res 65: 4827–4835.

Van 't Veer LJ, Dai H, Van de Vijver MJ, et al: 2002. Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530–536.

Warner S: 2004. Diagnostics + therapy = theranostics. Scientist 18:38.

Wu C, Ying H, Grinnell C, et al: 2007. Simultaneous targeting of multiple disease mediators by a dual-variable-domain immunoglobulin. Nat Biotechnol 25:1290–1297.

Zurita AJ, Troncoso P, Cardó-Vila M, et al: 2004. Combinatorial screenings in patients: the interleukin-11 receptor alpha as a candidate target in the progression of human prostate cancer. Cancer Res 64: 435–439.

Fibroblast Growth Factor 4 Gene Therapy for Chronic Ischemic Heart Disease

Navin K. Kapur and Jeffrey J. Rade*

Therapeutic myocardial angiogenesis and arteriogenesis represent a novel treatment strategy for patients with angina refractory to traditional medical and surgical therapies. The fibroblast growth factors are a family of proteins that are known mediators of angio-/arteriogenesis. Based on promising preclinical animal data, a series of four randomized placebo-controlled clinical trials have been conducted to determine the safety and efficacy of local delivery of fibroblast growth factor 4 with the use of adenovirus-vector-mediated gene transfer to induce myocardial angio-/arteriogenesis in patients with stable angina. This review describes the scientific rationale underlying these clinical trials, provides an overview of their results, and discusses the implications for future studies. (Trends Cardiovasc Med 2008;18:133–141) © 2008, Elsevier Inc.

• Introduction

Nearly nine million people in the United States have chronic stable angina pectoris caused by flow-limiting coronary atherosclerosis (Rosamond et al. 2007). Mechanical revascularization with percutaneous coronary angioplasty/stenting or coronary artery bypass surgery is largely effective at relieving angina when medications alone fail to control symptoms. Unfortunately, many patients with chronic angina are anatomically unsuitable for mechanical revascularization or continue to have a significant degree of ischemia despite these procedures.

The coronary microcirculation is a vast plexus of interconnecting arterioles and capillaries, although under normal conditions there is little collateral blood flow between epicardial vascular territories. As atherosclerosis progressively limits antegrade epicardial blood flow, changes in hemodynamic forces initiate a complex process of microvascular remodeling that promotes the development of collateral circulation (Schaper and Ito 1996). Although the term angiogenesis is frequently used to refer to the global process of collateral blood vessel formation, it more precisely refers to the sprouting of new capillary networks (Carmeliet 2000). Arteriogenesis refers to the process of investing either nascent or preexisting capillary channels with smooth muscle cells to form conduit arterioles. Of these two interrelated but distinct processes, arteriogenesis is likely the more critical in terms of being able to meaningfully augment collateral blood flow to ischemic myocardium subtended by occluded first- and second-order epicardial vessels. Nearly all patients with chronic angina form collateral vessels to some variable degree, with those >100 μm in diameter being visible to the naked eye during coronary angiography (Figure 1). Although collateral blood flow is often sufficient to prevent ischemia at rest, in approximately two-thirds of patients, there is insufficient flow augmentation to prevent angina during exercise (Pohl et al. 2001). The goal of therapeutic angio-/arteriogenesis is to relieve effort-induced angina pectoris by increasing blood flow to areas of ischemic myocardium via the development of new or the enhancement of existing coronary collateral blood vessels.

• Biology of Fibroblast Growth Factor 4

Angiogenesis and arteriogenesis involve a complex interplay between an array of different growth factors and target effector cells. One class of growth factors that is as an important mediator of these processes is the fibroblast growth factor (FGF) family. The FGF family comprises 22 distinct polypeptide growth factors composed of 150 to 250 amino acids that share 15% to 65% homology (Itoh 2007). Many of the FGFs bind anionic glycosaminoglycans, such as heparan and heparin sulfate, and are therefore referred to as heparin-binding growth factors. Heparan sulfate expressed on cellular surfaces facilitates the binding of FGFs to one of four high-affinity receptors (FGFR1-4) with tyrosine kinase activity. Binding of FGFs to their receptors can stimulate a wide array of signaling pathways (Dailey et al. 2005). Differential control of signal

Navin K. Kapur and Jeffrey J. Rade are at the Division of Cardiology, Johns Hopkins School of Medicine, Baltimore, MD 21287, USA.

* Address correspondence to: Jeffrey J. Rade, MD, Division of Cardiology, johns Hopkins School of Medicine, Carnegie 568, 600N. Wolfe St., Baltimore, MD 21287, USA. Tel.: (+1) 410 614 1361; fax: (+1) 410 614 0096; e-mail: jjrade@jhmi.edu.

Epidermal Growth Factor Receptor Abnormalities in the Pathogenesis and Progression of Lung Adenocarcinomas

Ximing Tang,[1] Marileila Varella-Garcia,[3] Ana C. Xavier,[3] Erminia Massarelli,[1] Natalie Ozburn,[1] Cesar Moran[2] and Ignacio I. Wistuba[1,2]

Abstract   To identify the characteristics and sequence of epidermal growth factor receptor (EGFR) abnormalities relevant to the pathogenesis and progression of lung adenocarcinoma, we performed a precise mapping analysis of *EGFR* mutation, gene copy number, and total and phosphorylated EGFR protein expression for the same tissue sites. We examined normal bronchial and bronchiolar epithelium (NBE) and tumor tissues obtained from 50   formalin-fixed lung adenocarcinomas, including 24 *EGFR*-mutant primary tumors with nine corresponding lymph node metastases and 26 wild-type primary tumors. NBE in 12 of 24 (50%) mutant and 3 of 26 (12%) wild-type tumors harbored *EGFR* mutations; these NBE also showed a lack of *EGFR* copy number increase and frequent EGFR (69%) and phosphorylated EGFR (33%) overexpression. *EGFR* mutation and protein overexpression were more frequent in NBE sites within tumors than in NBE sites adjacent to and distant from tumors, suggesting a localized field effect. Sites with high and low *EGFR* copy numbers were heterogeneously distributed in six of nine primary tumors and in one of eight metastases. EGFR protein overexpression was significantly higher in metastasis sites than in primary tumors. We conclude from our findings that *EGFR* mutations and protein overexpression are early phenomena in the pathogenesis of lung adenocarcinoma and that *EGFR* mutation precedes an increase in gene copy number. In *EGFR*-mutant adenocarcinoma metastases, the higher levels of EGFR overexpression and more homogeneously distributed high gene copy numbers suggest tumor progression. Our findings have important implications for the development of new strategies for targeted chemoprevention and therapy in lung adenocarcinoma using EGFR inhibitors.

Epidermal growth factor receptor (EGFR), a tyrosine kinase (TK) member of the ErbB family, has shown frequent abnormalities in non–small cell lung carcinomas. These abnormalities include protein overexpression, gene amplification, and mutation (1–3). Somatic *EGFR* mutations have been identified in specific subsets of patients with lung adenocarcinoma, including never or light smokers, women, and patients of East Asian descent (4). The mutations cluster in the first four exons (18–21) of the TK domain of the gene, and ~90% of the mutations are composed of either an in-frame deletion in exon 19 or a specific missense mutation in exon 21 (4). An increase in *EGFR* gene copy number, including high polysomy and gene amplification shown by fluorescent *in situ* hybridization (FISH), has been detected in 22% of patients with surgically resected (stages I-IIIA) non–small cell lung carcinomas and correlated with EGFR protein overexpression (2). Higher frequencies (40-50%) of *EGFR* high copy number have been reported in patients with advanced non–small cell lung carcinoma   (5–10). Despite this knowledge, limited information is available on the role of EGFR abnormalities in the early pathogenesis and progression of lung adenocarcinomas.

Recently, we showed that mutation of the *EGFR* TK domain is an early event in the pathogenesis of lung adenocarcinoma and is detected in histologically normal bronchial and bronchiolar epithelium (NBE) in 43% of patients with *EGFR*-mutant tumors (11). We found that *EGFR* mutations were more frequent in normal epithelium within the tumor (43%) than in adjacent sites (24%), suggesting a localized field effect (11). However, no comprehensive information is available regarding the role of *EGFR* abnormalities, including gene mutation, increased copy number, and protein overexpression in the early pathogenesis and progression of lung adenocarcinomas.

Both *EGFR* gene mutations and high copy number (gene amplification and high polysomy identified by FISH) have been associated with sensitivity to the small-molecule TK inhibitors gefitinib and erlotinib in patients with lung

adenocarcinoma (5–18). However, some of these results have been rather controversial (9, 10, 19, 20). In these studies of gefitinib and erlotinib, most of the *EGFR* mutation and copy number analyses were done in very small tissue samples or in cytologic specimens obtained from primary tumor and metastasis sites in patients with advanced-stage lung cancer (5–9, 12–16). To date, no studies have been done to identify the characteristics of *EGFR* gene and protein expression abnormalities at different sites with respect to primary lung adenocarcinomas and in corresponding sites of metastasis, information that might resolve some of the controversy.

To identify the sequence of EGFR abnormalities involved in the pathogenesis and progression of lung adenocarcinoma, we did a precise mapping analysis correlating *EGFR* mutation, gene copy number, and protein expression in NBE fields, primary tumors, and corresponding lymph node metastases that were obtained from 50 patients with lung adenocarcinomas, including 24 patients with *EGFR*-mutant primary tumors with nine corresponding lymph node metastasis sites and 26 patients with *EGFR*–wild-type primary tumors.

Materials and Methods

Case selection

To map *EGFR* gene and protein expression abnormalities, we obtained formalin-fixed, paraffin-embedded lung adenocarcinoma tissue specimens from the Lung Cancer Specialized Program of Research Excellence Tissue Bank at The University of Texas M. D. Anderson Cancer Center (Houston, TX). The tumor tissue specimens came from 50 patients with surgically resected lung adenocarcinomas (tumor-node-metastasis stage I-IIIA) with known *EGFR* mutations in exons 18 to 21, as described previously (3, 11). This bank was approved by the M. D. Anderson Cancer Center Institutional Review Board.

Of these 50 patients, 24 patients had lung adenocarcinoma with *EGFR* mutations in exon 18 (*n* = 1), exon 19 (*n* = 13), and exon 21 (*n* = 10), and 26 patients had *EGFR*–wild-type lung adenocarcinoma. The patients' clinicopathologic features are summarized in Table 1. All lung adenocarcinomas were of mixed histologic subtype (WHO classification; ref. 21). None of the patients had received cytotoxic and/or targeted therapy. Clinical staging was based on the revised International System for Staging Lung Cancer (22).

EGFR abnormality mapping

We retrospectively reviewed H&E-stained histology sections of primary tumor, lymph node metastases, and adjacent normal lung tissue specimens to identify tissue foci available for EGFR abnormality analyses. The EGFR abnormalities included *EGFR* mutations in exons 18 and 21, as shown by microdissection and PCR-based sequencing; *EGFR* copy number, as shown by FISH; and total EGFR and phosphorylated EGFR (pEGFR), as shown by immunohistochemical analyses. Representative examples of these molecular changes are illustrated in Fig. 1.

We used serial 5-μm-thick histology sections for the tissue microdissection, FISH, and immunohistochemical analyses. We identified a total of 316 noncontiguous tumor and epithelial foci from among 142 NBE specimens (obtained from 50 patients; 2.84 sites/patient), 144 primary tumors (from 50 patients; 2.88 sites/patient), and 30 lymph node metastases (from 9 patients; 3.3 sites/patient). We examined NBE and primary tumors in both *EGFR*-mutant and *EGFR*–wild-type cases and metastasis sites in *EGFR*-mutant cases only. All epithelial foci consisted of normal or mildly hyperplastic epithelia that harbored small bronchi (65 sites) and bronchioles (77 sites).

The NBE specimens were obtained from three different locations based on their relationship to the tumors: within the tumor (47 sites), ≤5 mm from the tumor margin (adjacent to tumor; 63 sites), and

Table 1. Clinicopathologic features of patients with lung adenocarcinomas examined for *EGFR* abnormalities in tumors and adjacent normal epithelium

Features/samples	*EGFR* status		
	Mutant (*n* = 24)	**Wild-type** (*n* = 26)	**Total** (*n* = 50)
Mean age (y)	61.3	62.7	62.1
Gender			
Female	19 (79%)	13 (50%)	32
Male	5 (21%)	13 (50%)	18
Ethnicity			
East Asian	13 (54%)	9 (35%)	22
Not East Asian	11 (56%)	17 (65%)	28
Smoking history			
Never	16 (67%)	9 (35%)	25
Former	7 (29%)	10 (38%)	17
Current	1 (4%)	7 (27%)	8
Stage of disease			
I	11 (46%)	15 (58%)	26
II	5 (21%)	4 (15%)	9
IIIA	8 (33%)	7 (27%)	15

>5 mm from the tumor margin ("distant" lung; 32 sites). We did not detect squamous metaplastic or dysplastic lesions in the bronchial structures or atypical adenomatous hyperplasias in the alveolar tissue. We identified small bronchi on the basis of well-defined smooth muscle and discontinuous cartilage layers. Bronchioles were defined as small conducting airways lacking well-defined smooth muscle wall or cartilage layers. We assessed the location of the small bronchial and bronchiolar respiratory epithelium examined for *EGFR* abnormalities based on the epithelia's location in relation to the tumor tissue in the corresponding histology sections, as previously described (11).

Microdissection and DNA extraction

Approximately 1,000 cells were precisely microdissected from 8-μm-thick, H&E-stained, formalin-fixed, paraffin-embedded histology sections for each site using laser capture microdissection (Arcturus Engineering Laser Capture Microdissection System; MDS Analytical Technologies), as previously described (11). To prevent the nonspecific binding of the mutant cells to the microdissection cap film, the microdissected tissue samples were redissected from the film under stereomicroscope visualization using fine needles (25-gauge 5/8-inch needles). We then extracted the DNA using 25 μL of PicoPure DNA Extraction solution containing proteinase K and incubated the DNA at 65°C for 20 h. Subsequently, proteinase K was inactivated by heating samples at 95°C for 10 min.

EGFR mutation analysis

Exons 18 and 21 of *EGFR* were PCR-amplified using DNA extracted from microdissected NBE and tumor cells, as previously described (3, 11). Each PCR was done using HotStarTaq Master Mix (Qiagen) for 40 cycles at 94°C for 30 s, 63°C for 30 s, and 72°C for 30 s, followed by a 7-min extension at 72°C. PCR products were directly sequenced using the Applied Biosystems PRISM dye terminator cycle sequencing method (Perkin-Elmer Corp.). We confirmed all sequence variants by independent PCR amplifications from at least two independent microdissections and sequenced the variants in both directions.

EGFR FISH analysis

We analyzed the gene copy number per cell using the LSI EGFR SpectrumOrange/CEP 7 SpectrumGreen Probe (Abbott Molecular), as previously described (5). Histology sections were incubated at 56°C overnight and deparaffinized by washing in CitriSolv (Fisher Scientific). After incubation in 2× SSC buffer (pH 7.0) at 75°C for 15 to 25 min, the histology sections were digested with proteinase K (0.25 mg/mL in 2× SSC) at 37°C for 15 to 25 min, rinsed in 2× SSC (pH 7.0) at room temperature for 5 min, and dehydrated using ethanol in a series of increasing concentrations (70%, 85%, 100%). We applied the EGFR SpectrumOrange/CEP 7/SpectrumGreen probe set (Abbott Molecular) onto the selected area, according to the manufacturer's instructions, on the basis of the tumor foci seen on each slide. We then covered the hybridization area with a glass coverslip and sealed the coverslip with rubber cement. The slides were incubated at 80°C for 10 min for codenaturation of chromosomal and probe DNA and then placed in a humidified chamber at 37°C for 20 to 24 h to allow hybridization to occur. Posthybridization washes were done in 1.5 mol/L of urea and 0.1× SSC (pH 7.0-7.5) at 45°C for 30 min and in 2× SSC for 2 min at room temperature. After the samples were dehydrated in a series of increasing ethanol concentrations, 4′,6′-diamidino-2-phenylindole (0.15 mg/mL in Vectashield Mounting Medium; Vector Laboratories) was applied for chromatin counterstaining. FISH analysis was done independently by two authors (M. Varella-Garcia and A. C. Xavier), who were blinded to the patients' clinical characteristics and all other molecular variables. Patients were classified into six FISH strata according to the frequency of cells with the EGFR gene copy number and referred to the chromosome 7 centromere, as follows: (a) disomy (≥3 copies in <10% of cells); (b) low trisomy (3 copies in 10% to 40% of the cells, ≥4 copies in <10% of cells); (c) high trisomy (3 copies in ≥40% of cells, ≥4 copies in <10% of cells); (d) low polysomy (≥4 copies in 10–40% of cells); (e) high polysomy (≥4 copies in ≥40% of cells); and (f) gene amplification (ratio of EGFR gene to chromosome ≥2, presence of tight EGFR gene clusters and 15 copies of EGFR per cell in 10% of the analyzed cells). The high polysomy and gene amplification categories were considered to be high EGFR copy number, and the other categories were considered to be nonincreased EGFR copy number, as previously published (5). Analysis was done in approximately 50 nuclei per tumor and epithelial site, and the section of the area was guided by image captured in the H&E-stained section.

Immunohistochemical staining

Tissue histology sections for immunohistochemical analyses were deparaffinized, hydrated, heated in a steamer for 10 min with 10 mmol/L of sodium citrate (pH 6.0) for antigen retrieval, and washed in Tris buffer. Peroxide blocking was done with 3% H_2O_2 in methanol at room temperature for 15 min, followed by 10% bovine serum albumin in TBS with Tween 20 for 30 min at room temperature. For the EGFR analysis, tissue sections were incubated for 2 h with primary antibodies against the EGFR clone 31G7 (1:100 dilution; Zymed) and pEGFR Tyr 1086 (1:100 dilution; Invitrogen). Tissue sections were then incubated for 30 min with the secondary antibody (EnVision+ Dual Link labeled polymer; DAKO), after which diaminobenzidine chromogen was applied for 5 min. The slides were then
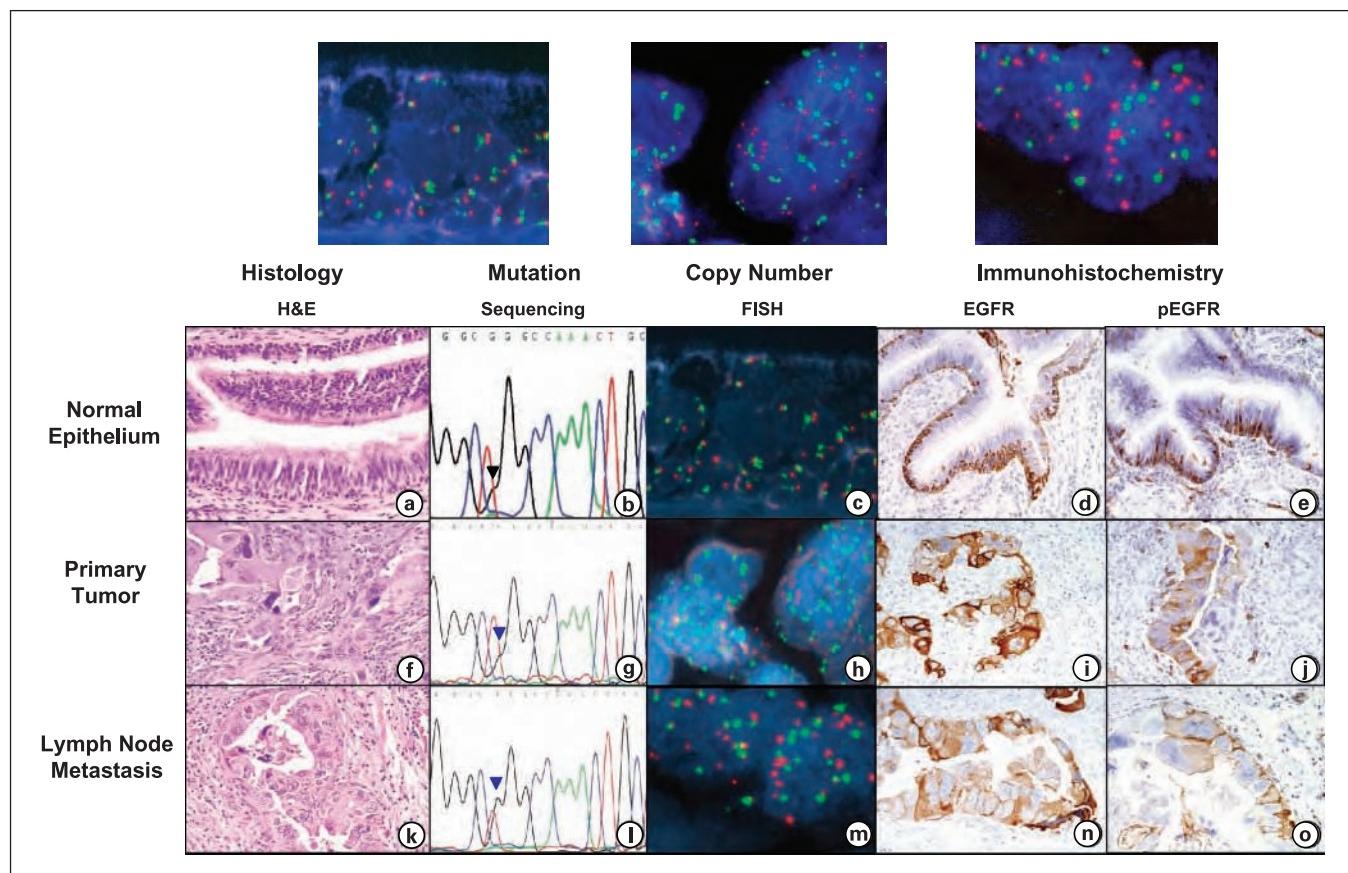
Fig. 1. A representative case of EGFR-mutant lung adenocarcinoma: EGFR gene and protein expression abnormalities in NBE (A–E), primary tumor (F–J), and lymph node metastasis (K–O) sites. Histologic characteristics (A, F, and K) of tissue sections stained with H&E (magnification, ×100). PCR-based EGFR sequencing (B, G and L) of the same EGFR mutation in exon 21 (L858R, black arrowhead) in NBE (B), primary tumor (G), and lymph node metastasis sites (L). EGFR FISH analysis (C, H and M) of low trisomy (low copy number) in the NBE sample (C), high polysomy in the primary tumor site (H), and gene amplification (M) in the metastasis site. Immunohistochemical analysis (D, I, N, E, J, and O) of high EGFR and pEGFR expression in the membrane and cytoplasm in all three types of samples.

Table 2. Frequency of *EGFR* gene mutation and protein overexpression in histologically normal bronchial and bronchiolar epithelium obtained from *EGFR*-mutant and wild-type lung adenocarcinomas

EGFR abnormality in NBE	Cases			Sites		
	Mutant	**Wild-type**	**Total**	**Mutant**	**Wild-type**	**Total**
Mutation by sequencing						
Number	24	26	50	85	57	142
Mutant	12 (50%)*	3 (12%)*	15 (30%)	22 (26%)	8 (14%)	30 (21%)
Protein overexpression by immunohistochemistry[†]						
Number	23	26	49	78	56	134
EGFR	19 (83%)	15 (58%)	34 (69%)	52 (67%)	35 (63%)	87 (65%)
pEGFR	10 (44%)	6 (23%)	16 (33%)	24 (31%)	12 (21%)	36 (27%)

*$P = 0.003$.
[†]Positive immunohistochemical overexpression score >200 (range, 0-400).

counterstained with hematoxylin and topped with a coverslip. For EGFR and pEGFR expression, antibody specificity was confirmed using blocking peptide and phosphatase incubation experiments. For the control experiments, we used formalin-fixed and paraffin-embedded pellets from lung cancer cell lines with confirmed EGFR and pEGFR overexpression. Thyroid transcription factor-1 (TITF-1) antibody (1:100 dilution, Cell Marque) was used for the identification of TITF-1–positive cells. All four antibodies were incubated for 1.5 h at room temperature. Immunohistochemistry results were scored jointly by two authors (X. Tang and I.I. Wistuba), who were blinded to clinical and other molecular variables. Immunostaining of the cell membrane and cytoplasm for EGFR and pEGFR was evaluated by light microscopy (magnification, ×20). A semiquantitative approach was used to generate a score for each tissue site, as previously described (2, 23, 24). Membrane and cytoplasm stains were recorded separately. We defined the intensity score as follows: 0, no appreciable staining in the NBE or malignant cells; 1, barely detect-

able staining in NBE or malignant cells compared with the stromal elements; 2, readily appreciable staining; 3, dark brown staining of cells; and 4, very strong staining of cells. The score was also based on the fraction of cells showing a given staining intensity (0-100%). We calculated the immunohistochemical scores by multiplying the intensity and extension, and the scores ranged from 0 to 400. For the statistical analyses, scores of 0 to 200 signified negative/low expression, and scores >200 indicated positive/overexpression, as previously reported (2, 23, 24). For the evaluation of nuclear TITF-1 immunohistochemical expression, 200 epithelial cells were quantified by light microscopy (magnification, ×20), and a score (range, 0-100) expressing the percentage of positive cells was obtained.

Statistical analysis

All relationships between categorical variables were assessed using χ^2 and Fisher's exact tests. $P < 0.05$ values were considered statistically significant.

Table 3. *EGFR* mutation and protein overexpression in histologically normal epithelium by location

| *EGFR* abnormality in NBE	Location in relation to the tumor			Structure	
	Inside	**Adjacent**	**Distant**	**Bronchiole**	**Small bronchus**
Mutation					
Mutant tumor	11/31 (36%)*	10/35 (29%)	1/17 (6%)*	10/43 (23%)	12/42 (29%)
Wild-type tumor	2/15 (13%)	3/28 (11%)	1/15 (7%)	4/34 (12%)	2/23 (9%)
All tumors	13/46 (28%)	13/63 (21%)	2/32 (6%)	14/77 (19%)	14/65 (22%)
EGFR overexpression[†]					
Mutant tumor	24/29 (83%)[‡]	20/33 (61%)[‡]	8/16 (50%)[‡]	18/38 (47%)[§]	34/40 (85%)[§]
Wild-type tumor	10/15 (67%)	17/28 (61%)	8/13 (62%)	14/33 (42%)[§]	21/23 (91%)[§]
All tumors	34/44 (77%)	37/61 (61%)	16/29 (55%)	32/71 (45%)[§]	55/63 (87%)[§]
pEGFR overexpression[†]					
Mutant tumor	13/29 (45%)[‖]	5/33 (15%)[‖]	6/16 (38%)[‖]	10/38 (26%)	14/40 (35%)
Wild-type tumor	5/15 (33%)	5/28 (18%)	2/13 (15%)	2/33 (6%)[§]	10/23 (44%)[§]
All tumors	18/44 (41%)	10/61 (16%)	8/29 (28%)	12/71 (17%)[¶]	24/63 (38%)[¶]

*Comparison of NBE from inside tumor vs. NBE distant ($P = 0.02$).
[†]Positive immunohistochemical overexpression score >200 (range 0-400).
[‡]Comparison of NBE from inside tumor vs. NBE adjacent + distant ($P = 0.02$)
[§]Comparison of NBE from bronchiole vs. small bronchus ($P < 0.001$).
[‖]Comparison of NBE from inside tumor vs. NBE adjacent + distant ($P = 0.038$).
[¶]Comparison of NBE from bronchiole vs. small bronchus ($P = 0.006$).

Results

EGFR abnormalities in the early pathogenesis of lung adenocarcinomas

Patterns of EGFR mutation in NBE. We previously reported our finding of mutations in exons 19 and 21 of *EGFR* in at least one site of microdissected NBE obtained from lung cancer specimens from 9 of 21 (43%) patients with *EGFR*-mutant adenocarcinomas, with no such mutations found in any of 26 respiratory epithelium foci from 16 patients with wild-type tumors (11). In the present study, using the same methodol-ogy, we analyzed for *EGFR* mutation in NBE obtained from an additional 3 patients with an *EGFR*-mutant and 10 patients with *EGFR*–wild-type lung adenocarcinomas. Combining both data sets, the overall rate of mutation in NBE from *EGFR*-mutant tumors was 50%. In the wild-type tumor cases, we detected *EGFR* exon 19 deletions (15 bp, 746-750) in six sites of small bronchial (*n* = 4) and bronchiolar (*n* = 2) NBE obtained from three wild-type tumors (Table 2). Thus, an *EGFR* mutation was found in NBE in 3 of 26 (12%) wild-type adenocarcinomas and in 8 of 57 (14%) of the microdissected epithelial sites (Table 2).
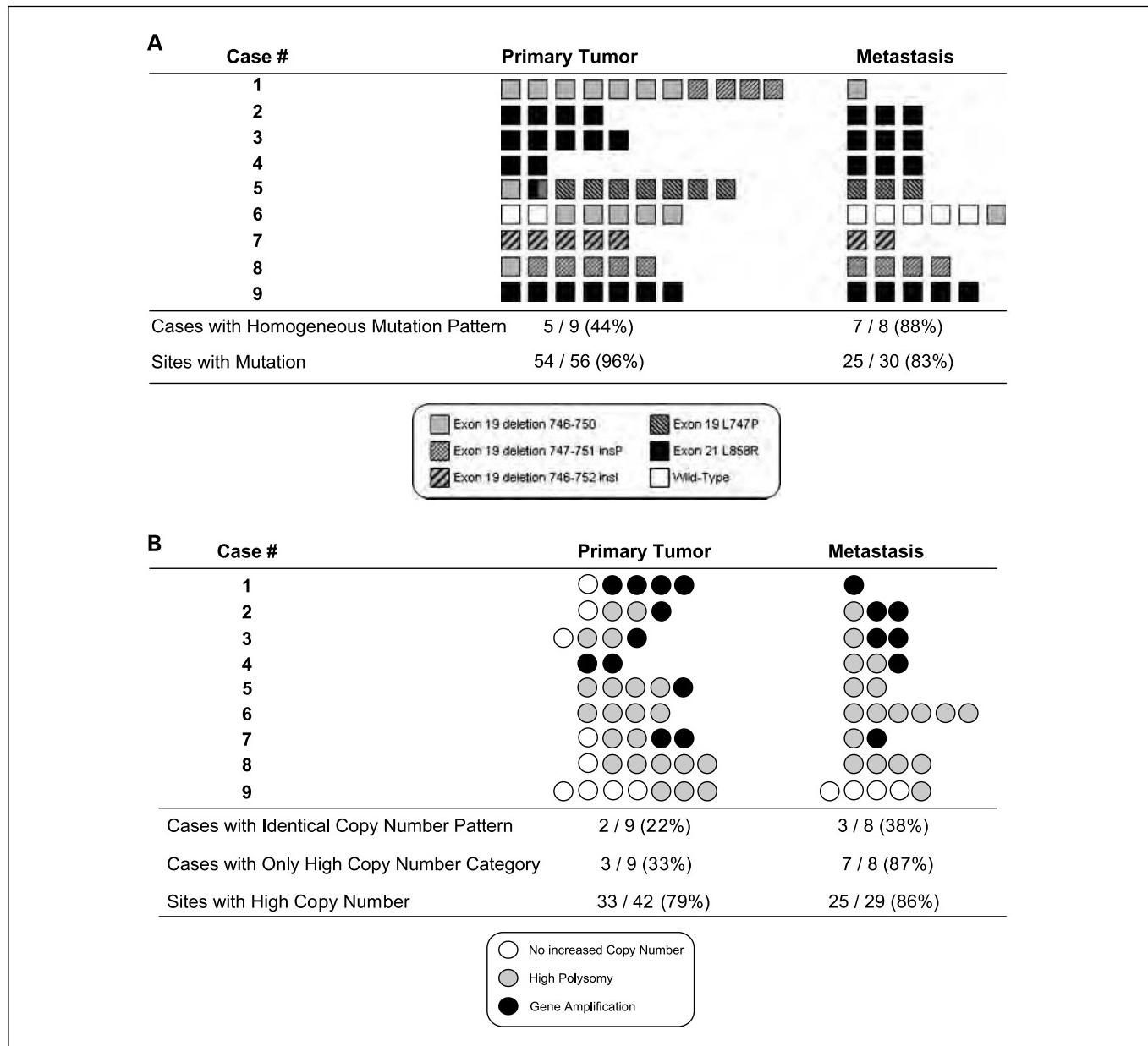
Fig. 2. *A*, *EGFR* mutation pattern in 56 primary tumor and 30 lymph node metastasis sites obtained from nine patients with *EGFR*-mutant lung adenocarcinomas. A homogeneous mutation pattern was detected in five primary tumors (*cases 2, 3, 4, 7*, and *9*) and all but one (*case 6*) metastasis case. *Case 6*, mixed wild-type and mutant sites in both primary tumor sites and corresponding metastases. *B*, *EGFR* copy number pattern shown by FISH in 42 primary tumor and 29 lymph node metastasis sites obtained from nine patients with *EGFR*-mutant lung adenocarcinomas. Different FISH copy number categories (low vs. high) were found in six of nine primary tumors and in one of eight corresponding metastases. Positive *EGFR* FISH expression included high polysomy and gene amplification, and negative *EGFR* FISH expression included disomy and trisomy.

Table 4. Summary of *EGFR* abnormalities by sites in nine primary lung adenocarcinomas and corresponding lymph node metastases

EGFR abnormality/ number of sites	Primary tumor	Metastases
Mutation		
Number of sites examined	56	30
Mutation positive	54 (96%)*	25 (83%)*
Copy no.		
Number of sites examined	42	29
Low copy no.	9 (21%)	4 (14%)
High copy no.	33 (79%)	25 (86%)
High polysomy	22 (52%)	18 (62%)
Gene amplification	11 (26%)	7 (24%)
Protein overexpression[†]		
Number of sites examined	65	31
EGFR	42 (65%)[‡]	30 (97%)[‡]
pEGFR	9 (14%)[§]	21 (68%)[§]

*The same case harbored two primary tumor and five metastasis sites with *EGFR*-wild-type sequence.

[†]Positive immunohistochemical expression score >200 (range 0-400).

[‡]Primary tumor vs. metastasis (P = 0.02).

[§]Primary tumor vs. metastasis (P = 0.00001).

The combined data showed that NBE with mutant *EGFR* was detected in the small bronchi (13 of 64, 20%) and bronchioles (17 of 78, 22%) of both mutant and wild-type tumor cases. Overall, however, the mutation frequency was higher in NBE samples microdissected from within the tumor (13 of 47, 28%) than in samples obtained from adjacent tissue and tissue distant from the tumors (17 of 95, 18%; Table 3).

In our previously reported comparison of NBE and corresponding tumors (16 specimens), we always observed identical *EGFR* mutations in both sites examined (11). In this study, we have expanded the number of NBE sites (n = 85) examined for the mutation in patients with *EGFR*-mutant adenocarcinomas and detected five sites (6%) from three cases in which NBE showed mutations different from the ones detected in the corresponding tumor specimens (data not shown). Importantly, in all cases with a mutation in NBE, an identical mutation was detected in at least one site of the corresponding tumor specimen. Thus, in this expansion of our previous study (11), a relatively more heterogeneous *EGFR* mutation pattern of the respiratory field was detected in NBE microdissected from mutant lung adenocarcinomas, but most NBE and corresponding tumors shared the same mutation.

EGFR copy number and correlation with gene mutation in NBE. To determine the morphologic stage at which *EGFR* copy abnormalities arise in *EGFR*-mutant adenocarcinomas, we did a precise mapping analysis and examined *EGFR* copy number in 21 NBE sites obtained from nine mutant adenocarcinomas using FISH. All nine tumor specimens showed at least one site with a high copy number. These epithelial sites were also examined in the *EGFR* mutation analysis. Most NBE (14 of 21, 67%) showed no *EGFR* FISH abnormalities (disomy), including four *EGFR*-mutant sites with exon 19 (15 bp) deletions and exon 21 (L858R) point mutations. Trisomy was detected in seven (33%) NBE sites obtained from six (67%) cases. We did not identify any NBE with *EGFR* amplifications or a high level of polysomy, which have been defined as high gene copy number. In contrast, the nine tumors mapped showed significantly higher frequency of *EGFR* amplification (11 of 42 sites, 26%; $P < 0.018$) or a high level of polysomy (22 of 42, 52%; $P < 0.001$) compared with NBE. Our findings indicate that high *EGFR* copy number does not occur in peripheral NBE in *EGFR*-mutant lung adenocarcinomas and that gene mutations precede copy number abnormalities in the sequential pathogenesis of these tumors.

EGFR immunohistochemical expression and correlation with gene mutation in NBE. We evaluated the level of EGFR and pEGFR protein expression in 134 NBEs obtained from *EGFR*-mutant and wild-type lung adenocarcinomas. Overall, a

Fig. 3. Proposed sequence of *EGFR* abnormalities occurring in the early pathogenesis and progression of *EGFR*-mutant lung adenocarcinomas. NBE field, primary tumor, and metastasis sites. *Small circles*, NBE, which acquires *EGFR* mutations and EGFR protein (total and phosphorylated) overexpression (*gray circles*). In the primary tumor stage, the *EGFR* copy number increases (high polysomy and gene amplification) in small tumor foci (*striped ovals*). In the metastasis site, tumor cells show both *EGFR* mutation and high copy number throughout most of the lesion.
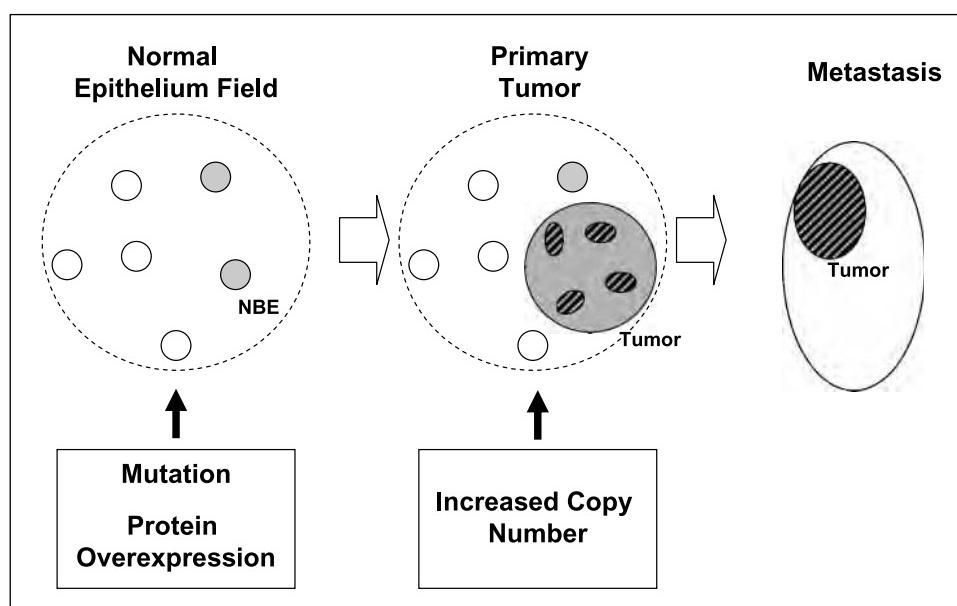
high level of EGFR (69%) and a moderate level of pEGFR (33%) expression were detected in NBE from patients with tumors (Table 2). However, EGFR and pEGFR were expressed to a greater degree in NBE sites obtained from patients with *EGFR*-mutant tumors than in patients with wild-type tumors (Table 2), although these differences were not statistically significant. The frequency of EGFR, but not of pEGFR, overexpression was higher in *EGFR*–wild-type NBE sites (85 of 111, 77%) than in mutant sites (14 of 24, 58%; $P = 0.039$). Of interest, NBE located inside tumors showed the highest frequency of EGFR and pEGFR overexpression compared with NBE located adjacent to and distant from tumors, especially in *EGFR*-mutant tumors (Table 3). Small bronchi also showed a higher frequency of overexpression of both markers compared with bronchioles (Table 3). Thus, the overexpression of EGFR and pEGFR is a common event in NBE from patients with lung adenocarcinomas, especially in *EGFR*-mutant tumors, and shows a localized field phenomenon effect similar to gene mutation.

TITF-1 immunohistochemical expression and* EGFR *mutation in NBE. Recently, on the basis of immunohistochemical findings of higher levels of nuclear TITF-1 expression, a crucial transcription factor of the lung, in *EGFR*-mutant lung adenocarcinomas compared with in wild-type tumors, it has been suggested that *EGFR*-mutant lung adenocarcinoma originates from the terminal respiratory unit (25), which is composed of alveolar cells and nonciliated bronchiolar epithelium. Its characteristics are highlighted by the expression of TITF-1 (25). We therefore investigated the correlation between *EGFR* mutation and TITF-1 nuclear expression in tumor and normal epithelium sites. *EGFR*-mutant lung adenocarcinomas (18 of 20 cases, 90%) showed higher expression of TITF-1 than did wild-type tumors (10 of 26 cases, 38%; $P < 0.001$). However, in immunohistochemical studies, we did not see a significant difference in the frequency of TITF-1 expression between *EGFR*-mutant (11 of 25 sites, 44%) and *EGFR*–wild-type (34 of 105 sites, 33%; $P = 0.273$) respiratory epithelia. Our findings therefore indicate that NBE cells expressing TITF-1 are not the exclusive precursors of *EGFR*-mutant adenocarcinomas. From these results, it is clear that these tumors do not originate exclusively from terminal respiratory unit structures.

EGFR abnormalities in the progression of lung adenocarcinomas

EGFR mutation pattern in primary tumors and corresponding metastasis. To identify the characteristics of EGFR abnormalities in the progression of mutant lung adenocarcinomas, we examined *EGFR* gene mutation, gene copy number, and protein expression in primary tumors and corresponding metastases by performing a detailed mapping analysis of tumor specimens. For this study, we selected nine lung adenocarcinomas with known *EGFR* mutations in exon 19 ($n = 5$) and exon 21 ($n = 4$), and with lymph node metastases for which there was sufficient tissue to perform our mapping analysis.

For the mutation analysis of *EGFR* exons 19 and 21, we did precise tissue microdissection from noncontiguous primary tumor foci ($n = 56$ sites, 6.2 sites/tumor; range 2-11 sites) containing at least 1,000 cells. Surprisingly, four of the nine primary tumorsexamined showed mixed *EGFR* gene patterns (Fig. 2A): three showed two or more types of mutations, and one showed five sites with exon 19 (15 bp, 746-750) dele-

tion and two sites with the wild-type *EGFR* gene. *EGFR* mutation analysis of 30 corresponding lymph node metastasis sites from the nine *EGFR*-mutant cases (3.3 sites/case; range 1-6 sites) detected only one type of *EGFR* mutation in all tumor sites in each case, and the mutation was always present in at least one site of the corresponding primary tumor. Similar to the corresponding primary tumor, one metastasis case showed *EGFR*–wild-type (five sites) and *EGFR*-mutant [one site, exon 19 (15 bp, 746-750) deletion] tumor sites (Fig. 2A). All these findings were confirmed by sequencing analyses of independently microdissected samples. In summary, our findings showed a relatively high level of heterogeneity for the *EGFR* mutation, and several tumor cell clones had mutation patterns in the primary tumor specimens that differed from the mutation patterns in the lymph node metastasis sites.

EGFR copy number abnormalities in primary tumors and corresponding metastasis. We used FISH to investigate the *EGFR* gene copy number abnormalities in 42 primary tumor sites (2.1 sites/case; range 2-7 sites) and 29 metastasis sites (3.2sites/case; range 1-6 sites), which were also examined for the mutation analysis. Overall, all primary tumors and corresponding metastases showed at least one site of high gene copy number (high polysomy or gene amplification; Fig. 2B). However, six (67%) primary tumor cases and one (11%) metastasis case showed at least one site without high copy number (disomy in one primary tumor site, high trisomy in one metastasis site, and low polysomy in seven primary and three metastasis sites; Fig. 2B). Thus, *EGFR* copy number heterogeneity was higher in primary tumor sites than in corresponding metastasis sites.

EGFR immunohistochemical expression in primary tumors and corresponding metastasis sites. In the nine *EGFR*-mutant lung adenocarcinoma cases mapped for *EGFR* abnormalities, we examined both primary tumors and the corresponding lymph node metastases for EGFR and pEGFR immunohistochemical expression. For both tumor locations combined, 96 distinct tumor sites were examined ($n = 65$ primary tumor sites, 7.2 sites/case; and $n = 31$ metastasis sites, 3.4 sites/case). Significantly higher levels of EGFR and pEGFR expression were detected in metastasis sites compared with primary tumor sites (Table 4). No correlation between EGFR and pEGFR expression and *EGFR* copy number status by FISH was detected.

Discussion

Using a detailed molecular pathology mapping strategy, we determined the sequence of EGFR abnormalities in the early pathogenesis of *EGFR*-mutant lung adenocarcinomas and identified the pattern of EGFR changes in the progression of *EGFR*-mutant lung adenocarcinomas from primary tumors to lymph node metastasis. First, we showed that *EGFR* mutations precede gene copy number abnormalities in the pathogenesis of these tumors and that EGFR and pEGFR immunohistochemical protein expressions are frequent events in histologically normal peripheral bronchial and bronchiolar epithelium adjacent to lung adenocarcinomas. Second, our data indicated that although primary lung adenocarcinomas show some degree of *EGFR* gene copy number heterogeneity, this phenomenon is rare in metastases. Although these findings can be considered tumor progression phenomena, they also have important clinical implications from the standpoint of making decisions regarding the use of EGFR TK inhibitor therapy on the basis of the finding of *EGFR* gene abnormalities.

Despite the evidence showing that atypical adenomatous hyperplasia is a precursor of peripheral lung adenocarcinomas (26), there is consensus that the pathogenesis of most adeno- carcinomas is unknown. Our previously reported findings of an *EGFR* mutation in NBE in 9 of 21 (43%) patients with *EGFR*-mutant adenocarcinomas indicated that the *EGFR* gene mutation is an early event in the pathogenesis of lung adeno- carcinoma (11). In this study, we have investigated normal epithelium from additional patients with *EGFR*-mutant or wild-type lung adenocarcinomas and specifically have two new findings in this study; (*a*) we detected an *EGFR* mutation (exon 19, 15 bp deletion, 746-751) in six sites of small bronchial and bronchiolar epithelium obtained from three patients with wild-type adenocarcinoma, and (*b*) whereas an identical mu- tation was detected in the majority of specimens of mutant normal epithelium compared with the corresponding invasive tumor (75% of cases and 94% of sites), we found few normal epithelium sites (6%) in three of 12 cases (25%) of *EGFR*- mutant tumors, demonstrating the existence of a different mu- tation pattern between normal epithelium and the correspond- ing invasive tumor. All these data reinforce the concept of a field effect phenomenon in *EGFR* mutations in lung adenocar- cinoma pathogenesis that affects histologically normal bron- chial and bronchiolar respiratory epithelia.

We have previously shown that molecular abnormalities occur in a stepwise fashion in the sequential pathogenesis of squamous cell carcinoma of the lung, with molecular changes commencing in histologically normal bronchial epithelium in smokers and in patients with lung cancer (27, 28). Our find- ings suggest that *EGFR* abnormalities also occur sequentially in the early pathogenesis of lung adenocarcinoma, with a mu- tation commencing in histologically normal epithelium and a high *EGFR* copy number appearing at the invasive tumor stage. A recent report (29) of selective gene amplification of the shorter allele of the *EGFR* intron 1 polymorphism CA sim- ple sequence repeat 1, which is the allele more frequently mu- tated in tumors harboring an *EGFR* mutation, also suggests that mutations occur earlier than copy number abnormalities in the pathogenesis of lung adenocarcinoma. Our findings of frequent EGFR (69%) and pEGFR (33%) protein overexpres- sion in normal distal bronchial and bronchiolar epithelium from patients with either *EGFR*-mutant or *EGFR*–wild-type lung adenocarcinomas indicate a field phenomenon in the peripheral airway. A relatively high frequency of EGFR pro- tein expression has also been reported in centrally located, histologically normal (42%) and hyperplastic (54%) bronchial epithelium from smokers (23). In addition, our data indicate that the mechanisms of protein overexpression seem to be un- associated with high gene copy number and mutation in NBE. Other mechanisms can explain EGFR overexpression in nor- mal epithelial cells, including ligand-dependent up-regulation and activation, as well as inhibition of endocytosis-related protein down-regulation in the cell membrane (30).

Based on findings of higher levels of immunohistochem- ical expression of nuclear TITF-1, a crucial transcription fac- tor of the lung, in *EGFR*-mutant lung adenocarcinomas compared with wild-type tumors, it has been suggested that *EGFR*-mutant lung adenocarcinoma originates from the terminal respiratory unit (25). We found *EGFR* mutations in microdissected histologically normal epithelial cells from small bronchi and bronchioles, which supports the concept

of adenocarcinomas arising from the peripheral lung airway. Our findings indicate that NBE cells expressing TITF-1 are not the exclusive precursors of *EGFR*-mutant adenocarcino- mas. From this finding, it is clear that these tumors do not originate exclusively from terminal respiratory unit struc- tures. In addition, we cannot exclude the possibility that common stem or progenitor cells for both bronchial and bronchiolar epithelium bear *EGFR* mutations.

It has been suggested that activating TK *EGFR* mutations are a potent oncogenic event by which mutant tumor cells become physiologically dependent on the continued activity of the phosphorylated protein for the maintenance of their malignant phenotype (31). Our detailed mapping analysis of the *EGFR* gene mutation and copy number of multiple precisely microdissected sites in nine mutant primary tu- mors and corresponding lymph node metastases showed an identical or monoclonal pattern of mutation in most (*n* = 5) primary tumors and all metastases. These findings corroborate the monoclonal concept of tumor development and the monoclonal evolution of metastases (32, 33). How- ever, two primary tumors lacking identical or monoclonal *EGFR*-mutant patterns harbored different sizes of exon 19 deletions (12 versus 15 bp and 15 versus 18 bp deletions). This finding could be explained by a tumor progression phe- nomenon in which the deletion size changed during the evo- lution of the malignancy. However, two very interesting primary tumors in our study exhibited findings that chal- lenged the concept of the monoclonal evolution of tumors. One case showed a single site with an exon 19 (15 bp) dele- tion, whereas the remaining eight sites lacked the deletion but showed a point mutation (TTA747CCA) in the same exon. Of interest, the three metastasis sites examined harbored the most frequent mutation detected in the primary lung tumor. The other case showed areas of wild-type and mutant *EGFR* in both primary tumors and metastases, a phenomenon that is difficult to explain and suggests that molecular events other than an *EGFR* mutation may be responsible for tumor development in lung adenocarcinomas. These findings were confirmed by the sequencing of independently microdis- sected samples. In the latter case, the finding of a high *EGFR* copy number (high polysomy) in wild-type tumor sites raises the possibility of an alternative explanation—that the wild- type allele is preferentially amplified in some tumor cells. As a result, the mutant allele is underrepresented and is not detectable by our current sequencing methodology.

Retrospective studies have provided data suggesting that a high *EGFR* gene copy number shown by FISH is associated with treatment response, time to progression, and survival in patients with advanced non–small cell lung carcinoma trea- ted with EGFR TK inhibitors (5-7, 10, 17). In these studies, high *EGFR* copy number as shown by FISH was defined as true gene amplification or high polysomy with equal to or more than four *EGFR* copies in ≥40% of cells (5, 34). Our mapping analysis of primary tumors and corresponding lymph node metastases in which we used the same *EGFR* FISH criteria showed that a frequent high copy number in mutant tumors was the most frequent pattern detected. De- spite the fact that most primary tumor sites and nearly all metastasis sites showed high copy numbers, high polysomy and gene amplification were heterogeneously distributed in both tumor locations. More importantly, five of nine (56%)

primary tumors and one metastasis (13%) showed one or more sites without an increased copy number (FISH negative). Similarly, EGFR and pEGFR immunohistochemical expression was less heterogeneous in primary tumors and more frequent in metastases. Taken together, these data suggest that *EGFR* copy number analyzed by FISH and protein expression analyzed by immunohistochemistry in small core biopsy or fine-needle aspiration specimens obtained from primary tumors, and more rarely from metastases, could miss these molecular changes, especially if only a small number of malignant cells are available for examination. In addition, if the suggested presence of *EGFR* high copy number correlates with sensitivity to EGFR TK inhibitors (5–7, 17), it is likely that metastases will show a better response to therapy than will primary tumors. This is an important consideration, in light of the fact that the site of origin (primary versus metastasis) of the tumor specimen was not reported and factored into the biomarker analyses in any of the clinical trials testing the efficacy of EGFR TK inhibitors in pa-tients with advanced non–small cell lung carcinoma in whom *EGFR* copy number determined by FISH was examined as a predictor of response and prognosis (5–7). Our results show that a better understanding of the pattern of molecular abnormalities and their corresponding biomarker expression, including primary tumors and the frequent metastases seen for this tumor type, is important in lung cancer.

In summary, our data suggest that gene mutations and protein overexpression are the earliest phenomena in *EGFR*-mutant lung adenocarcinoma, occurring at the NBE stage, and that this is followed by the development of a focal increase in copy number at the tumor stage (Fig. 3). At the metastasis sites, however, all three abnormalities were more frequent than they were in the primary tumors and were homogeneously distributed throughout the malignant cells.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Scagliotti GV, Selvaggi G, Novello S, Hirsch FR. The biology of epidermal growth factor receptor in lung cancer. Clin Cancer Res 2004;10:4227s–32s.

2. Hirsch FR, Varella-Garcia M, Bunn PA, Jr., et al. Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. J Clin Oncol 2003;21:3798–807.

3. Shigematsu H, Lin L, Takahashi T, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 2005;97:339–46.

4. Shigematsu H, Gazdar AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. Int J Cancer 2006;118:257–62.

5. Cappuzzo F, Hirsch FR, Rossi E, et al. Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. J Natl Cancer Inst 2005;97:643–55.

6. Tsao MS, Sakurada A, Cutz JC, et al. Erlotinib in lung cancer—molecular and clinical predictors of outcome. N Engl J Med 2005;353:133–44.

7. Hirsch FR, Varella-Garcia M, McCoy J, et al. Increased epidermal growth factor receptor gene copy number detected by fluorescence *in situ* hybridization associates with increased sensitivity to gefitinib in patients with bronchioloalveolar carcinoma subtypes: a Southwest Oncology Group Study. J Clin Oncol 2005;23:6838–45.

8. Jackman DM, Holmes AJ, Lindeman N, et al. Response and resistance in a non-small-cell lung cancer patient with an epidermal growth factor receptor mutation and leptomeningeal metastases treated with high-dose gefitinib. J Clin Oncol 2006; 24:4517–20.

9. Massarelli E, Varella-Garcia M, Tang X, et al. KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. Clin Cancer Res 2007;13:2890–6.

10. Bunn PA, Jr., Dziadziuszko R, Varella-Garcia M, et al. Biological markers for non-small cell lung cancer patient selection for epidermal growth factor receptor tyrosine kinase inhibitor therapy. Clin Cancer Res 2006;12:3652–6.

11. Tang X, Shigematsu H, Bekele BN, et al. EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. Cancer Res 2005;65:7568–72.

12. Paez JG, Janne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science 2004;304: 1497–500.

13. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N Engl J Med 2004;350: 2129–39.

14. Mitsudomi T, Kosaka T, Endoh H, et al. Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non-small-cell lung cancer with postoperative recurrence. J Clin Oncol 2005;23: 2513–20.

15. Han SW, Kim TY, Hwang PG, et al. Predictive and prognostic impact of epidermal growth factor receptor mutation in non-small-cell lung cancer patients treated with gefitinib. J Clin Oncol 2005;23: 2493–501.

16. Taron M, Ichinose Y, Rosell R, et al. Activating mutations in the tyrosine kinase domain of the epidermal growth factor receptor are associated with improved survival in gefitinib-treated chemorefractory lung adenocarcinomas. Clin Cancer Res 2005; 11:5878–85.

17. Han SW, Kim TY, Jeon YK, et al. Optimization of patient selection for gefitinib in non-small cell lung cancer by combined analysis of epidermal growth factor receptor mutation, K-ras mutation, and Akt phosphorylation. Clin Cancer Res 2006; 12:2538–44.

18. Pugh TJ, Bebb G, Barclay L, et al. Correlations of EGFR mutations and increases in EGFR and HER2 copy number to gefitinib response in a retrospective analysis of lung cancer patients. BMC Cancer 2007;7:128.

19. Sequist LV, Haber DA, Lynch TJ. Epidermal growth factor receptor mutations in non-small cell lung cancer: predicting clinical response to kinase inhibitors. Clin Cancer Res 2005;11:5668–70.

20. Johnson BE, Janne PA. Selecting patients for epidermal growth factor receptor inhibitor treatment: a FISH story or a tale of mutations? J Clin Oncol 2005;23:6813–6.

21. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Tumours of the lung. In: W.D. Travis E, Brambilla HK, Muller-Hermelink, C.C. Harris, editors. Pathology and genetics: Tumours of the lung, pleura, thymus and heart. Lyon: IARC; 2004. p. 9-124.

22. Mountain CF. Revisions in the International System for Staging Lung Cancer. Chest 1997;111: 1710–7.

23. Merrick DT, Kittelson J, Winterhalder R, et al. Analysis of c-ErbB1/epidermal growth factor receptor and c-ErbB2/HER-2 expression in bronchial dysplasia: evaluation of potential targets for chemoprevention of lung cancer. Clin Cancer Res 2006;12:2281–8.

24. Tsao AS, Tang XM, Sabloff B, et al. Clinicopathologic characteristics of the EGFR gene mutation in non-small cell lung cancer. J Thorac Oncol 2006;1:231–9.

25. Yatabe Y, Kosaka T, Takahashi T, Mitsudomi T. EGFR mutation is specific for terminal respiratory unit type adenocarcinoma. Am J Surg Pathol 2005;29:633–9.

26. Westra WH. Early glandular neoplasia of the lung. Respir Med 2000;1:163–9.

27. Wistuba II, Behrens C, Milchgrub S, et al. Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. Oncogene 1999;18:643–50.

28. Wistuba II, Behrens C, Virmani AK, et al. High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. Cancer Res 2000;60:1949–60.

29. Nomura M, Shigematsu H, Li L, et al. Polymorphisms, mutations, and amplification of the EGFR gene in non-small cell lung cancers. PLoS Med 2007;4:e125.

30. Grandal MV, Madshus IH, Epidermal growth factor receptor and cancer: control of oncogenic signalling by endocytosis. J Cell Mol Med 2008 [Epub ahead of print].

31. Gazdar AF, Shigematsu H, Herz J, Minna JD. Mutations and addiction to EGFR: the Achilles 'heal' of lung cancers? Trends Mol Med 2004;10:481–6.

32. Fearon ER, Hamilton SR, Vogelstein B. Clonal analysis of human colorectal tumors. Science 1987;238:193–7.

33. Garcia SB, Novelli M, Wright NA. The clonal origin and clonal evolution of epithelial tumours. Int J Exp Pathol 2000;81:89–116.

34. Varella-Garcia M. Stratification of non-small cell lung cancer patients for therapy with epidermal growth factor receptor inhibitors: the EGFR fluorescence *in situ* hybridization assay. Diagn Pathol 2006;1:19.

The Tumor Suppressor LKB1 Regulates Lung Cancer Cell Polarity by Mediating cdc42 Recruitment and Activity

Shumin Zhang,[1] Katherine Schafer-Hales,[1] Fadlo R. Khuri,[1] Wei Zhou,[1] Paula M. Vertino,[2] and Adam I. Marcus[1]

Departments of [1]Hematology and Oncology and [2]Radiation Oncology, Winship Cancer Institute, Emory University, Atlanta, Georgia

Abstract

The tumor suppressor LKB1 is mutated in 30% of non–small cell lung cancer (NSCLC) tumors and cell lines and is proposed to be a key regulator of epithelial cell polarity; however, how LKB1 regulates cancer cell polarity is not known. The experiments described herein show for the first time that LKB1 is a dynamic, actin-associated protein that rapidly polarizes to the leading edge of motile cancer cells. LKB1 proves to be essential for NSCLC polarity, because LKB1 depletion results in classic cell polarity defects, such as aberrant Golgi positioning, reduced lamellipodia formation, and aberrant morphology. To probe how LKB1 regulates these events, we show that LKB1 colocalizes at the cellular leading edge with two key components of the polarity pathway — the small rho GTPase cdc42 and its downstream binding partner p21-activated kinase (PAK). Importantly, LKB1 functionality is required for cdc42 polarization to the leading edge, maintaining active cdc42 levels, and downstream PAK phosphorylation. To do this, LKB1 interacts only with active form of cdc42 and PAK, but not with inactive cdc42. Taken together, these results show that LKB1 is a critical mediator of the NSCLC polarity program in lung cancer cells through a novel LKB1-cdc42-PAK pathway. [Cancer Res 2008;68(3):740–8]

Introduction

Cell polarization is essential for a broad range of cellular processes, such as mitosis, morphogenesis, and motility. In most eukaryotic cells, this cell polarity program is regulated by a complex network of signaling molecules, cytoskeletal elements, and extracellular cues that ultimately create a functionally and spatially distinct polarized region within the cell (1). One master regulator of this network is the small rho GTPase cdc42 (2); like other Rho GTPases, cdc42 cycles between a GTP-bound active state and GDP-bound inactive state. In its active GTP-bound state, cdc42 regulates key events of cell polarity that include lamellipodia formation (protrusive structure at the cellular leading edge of motile cells), Golgi reorientation, centrosome reorientation, and tight junction formation (2). To do this, active cdc42 relies on a series of downstream effectors that includes its direct binding partner p21-activated kinase (PAK; refs. 3, 4). Upon binding of active cdc42 to PAK, phosphorylation of PAK occurs and cell polarization is triggered (5).

Defects in cell polarity have been linked to cancer progression. It is proposed that the disruption of cell polarity within the epithelial cell lining serves as an initiator for cancer cell invasion into the surrounding environment (6). One potential regulator of the cancer polarity program is LKB1 (also known as STK11; ref. 7). LKB1 is a serine/threonine kinase that contains two nuclear localization sequences, a central kinase domain and a C-terminal farnesylation motif (8). The LKB1 gene is located on chromosome 19p13.3 and produces a 3.1-kb transcript that is expressed in many adult and fetal tissues (7). The N-terminal and C-terminal noncatalytic regions of LKB1 share no relatedness to other proteins. LKB1 was recently shown to be both a lung cancer tumor suppressor that serves as a repressor for the mTOR pathway of biosynthesis (9) and as a regulator of normal epithelial cell polarity (10). Approximately, 30% of non–small cell lung cancer (NSCLC) cell lines and tumors harbor an LKB1 mutation (11, 12). These mutations are primarily point mutations that were either nonsense or frameshift mutations (11).

Importantly, a landmark study showed that LKB1 activation causes complete polarization of single intestinal cells, even in the absence of junctional cell-to-cell contacts, traditionally a prerequisite for polarization (13). Because several LKB1 mutants are incapable of polarizing normal epithelial cells (14), it is intriguing to consider the possibility that one consequence of LKB1 mutation is aberrant cancer cell polarity. Nevertheless, the molecular details of how LKB1 mediates these cell polarity events is not well understood.

Here, we wanted to determine if LKB1 functions in NSCLC polarity, and if so, how LKB1 regulates these events. To address this, we performed a comprehensive analysis of endogenous LKB1 function in NSCLC cell lines using a wounding model of cell polarity and motility. We now show that LKB1 behaves as a dynamic, actin-associated protein that rapidly polarizes to the leading edge of motile cells. LKB1 is essential for maintaining cancer cell polarity because LKB1 defects affect a variety of cell polarity events. To do this, LKB1 regulates the activity and recruitment of cdc42 via its association with the active form of cdc42 and its binding partner PAK. Based upon these studies, we propose a novel LKB1-cdc42-PAK pathway that oversees cancer polarity.

Materials and Methods

Cell culture. All cells were maintained at 37°C in a humidified chamber as previously described (15).

Wounding assays. Cells were grown to confluency on either plastic dishes or coverslips (for microscopy studies) and then wounded using a 1-μL to 10-μL pipette tip. Over time cells will polarize and migrate into the wound. Cells were then processed for immunofluorescence, Western blotting, or immunoprecipitation as described below.

Transfections. Lipofectamine 2000 (Invitrogen) was used to transfect cell lines according to the manufacturer's protocol. Transfections were performed 24 h before wound induction. Oligofectamine (Invitrogen) was

used for small interfering RNA (siRNA) transfection according to manufacturer's protocol. LKB1 siRNA was used at 200 nmol/L 72 h before wound induction.

Immunofluorescence and confocal imaging. Immunofluorescence was performed as described previously (16). Antibodies against LKB1 (Abcam), cdc42 (Cytoskeleton), PAK (Santa Cruz), phosphorylated PAK (Biosource), GM-130 (Golgi; Calbiochem), or actin-phalloidin 488 (Invitrogen) were incubated with cells at 4°C overnight. Either Alexa 488, 563, or 633 secondary antibodies (Invitrogen) were used at a dilution of 1:500 and were incubated for 1 h at room temperature. Nuclear staining was performed by incubating cells with 0.4 μmol/L 4′,6-diamidino-2-phenyl-indole (DAPI) to mounting slides. Cells were imaged on a Zeiss LSM 510 META as described in ref. (15). In all cases, either a 63× or 100× Zeiss Plan-Apo oil objective was used (numerical aperature of 1.3 and 1.4, respectively). To quantitate fluorescence intensity, grayscale images were first thresholded and subjected to mean intensity analysis, with a designated region of interest in Metamorph 6.2 (Molecular Devices). The lamellipodia was defined as the region within 5 μm from the leading edge of the cell. Colocalization analysis was performed by using the "percentage colocalization" feature in Metamorph, which compared the amount of LKb1 colocalized with the protein of interest (e.g., cdc42, PAK) This analysis measured the percentage of LKB1 associated with the protein of interest by comparing signal overlap in thresholded images on a pixel-by-pixel basis. All images have contrast expansion performed in Adobe Photoshop.

Golgi reorientation polarity assays. To image Golgi positioning cells were fixed at different time points postwounding and stained for Golgi, LKB1, and DAPI as described above. Wound edge cells were divided into three 120° regions, with one region facing the wound edge (Supplementary Fig. S2). All cells with the Golgi facing the wound front were scored positive. For each time point, at least 20 cells were examined. Cell orientation was measured as shown in Supplementary Fig. S2. To do this, the angle was calculated between a line dropped from the midpoint of the cellular edge facing the wound and a line tracing the wound front. For examples, cells aligned perpendicular to the leading edge had a nearly 90° orientation, whereas cells aligned parallel to the wound front had a 0° orientation.

Western blotting. Western blotting was performed as previously described (16). Protein concentrations were determined by the bicinchoninic acid protein assay kit (Pierce). Equal amounts of protein from whole-cell lysates were solubilized in SDS sample buffer and separated on SDS 12.5% polyacrylamide gels. The same primary antibodies were used as described above with the appropriate secondary horseradish peroxidase–conjugated antibody (1:2,000).

Cdc42 activation assay. Cdc42 activation assays was performed with the cdc42 activation assay kit from Cell Biolabs, Inc. (STA-402) and used according to the manufacturer's protocol. Cell lysates were prepared from confluent monolayers of LKB1 wild-type H1299 cells or cells transfected 72 h prior with LKB1 siRNA. All the cells were scratched 200 times with a multitip pipette. Cells were washed with cold PBS twice and lysed with 700 μL of lysis buffer [25 mmol/L HEPES (pH 7.5), 150 mmol/L NaCl, 1% NP40, 10 mmol/L MgCl, 1 mmol/L EDTA, and 2% glycerol]. For each time point, 750 μg of protein was incubated with PAK-PBD conjugated agarose beads. Samples were incubated for 3 h at 4°C, washed, centrifuged, resuspended in 40 μL of 2× reducing SDS-PAGE sample buffer, and processed for Western blotting with an anti-Cdc42 mouse monoclonal antibody (described above). Blots were reprobed with mouse anti-LKB1 antibody to visualize the status of the LKB1/PAK-PBD interaction over time.

Immunoprecipitation. Immunoprecipitation was performed with the Catch and Release Reversible Immunoprecipitation System (Upstate) and used according to the manufacturer's protocol. Cells were first seeded into 10 cm² dishes then transfected with myc-PAK, FLAG-LKB1, and GFP-tagged cdc42 plasmids as described above. The GFP-cdc42-T17N is a dominant negative inactive cdc42, whereas the GFP-cdc42-Q61L is a constitutively active cdc42. Cells were lysed 24 h after transfection at 70% confluency in 300 μL kit supplied lysis buffer. Cell lysates (500 μg) were incubated with rabbit IgG or antirabbit flag antibody with 10 μL affinity ligand into a bead-coated column. The column was incubated at 4°C for 1 h and washed, and beads were resuspended into 70 μL of 1× denaturing elution buffer containing βME (5%). The samples were boiled for 5 min, centrifuged, and processed for Western blotting.

Results

LKB1 polarizes to the leading edge and associates with actin in motile NSCLC cells. To investigate LKB1 function in NSCLC, we assessed endogenous LKB1 localization in H1703 NSCLC cells (LKB1 wild-type). Cell motility and polarity were induced with a wounding assay that generates a cell free gap bordered by a confluent monolayer of cells. In this assay (17), cells that border the wound will polarize, generate a lamellipodia, and move into the gap over time. In confluent cells (i.e., stationary cells not bordering the wound), LKB1 was found throughout the cytoplasm and nucleus (Fig. 1A); however, in cells bordering and migrating into the wound, LKB1 drastically relocalized to or near the plasma membrane, at the leading edge of the cell, facing the direction of movement (Fig. 1A and B). This pattern of LKB1 localization was observed in all wild-type LKB1 NSCLC cell lines tested, including H1299 and H520 NSCLC cells (Supplementary Fig. S1). Furthermore, a time course showed that LKB1 polarization to the leading edge was rapid and occurred within 5 min postwounding (Fig. 1C). Quantitative fluorescent intensity analysis confirms these observations and showed that the percentage of motile cells with polarized LKB1 significantly increased in motile cells within 5 min ($P < 0.05$), as well did LKB1 fluorescent intensity at or near the plasma membrane (Fig. 1D). To our knowledge, this is the first time endogenous LKB1 has been successfully visualized and shows a distinct and rapid polarization to the leading edge of motile NSCLC cells.

The actin cytoskeleton is one of the primary components governing cell motility and polarity and is enriched in the lamellipodia at the leading edge; therefore, we next determined whether LKB1, which polarizes to the leading edge (Fig. 1), is associated with actin. To do this, cells migrating into the wound were fixed and costained for LKB1 and actin. LKB1 showed a highly significant colocalization with actin only at the leading edge of motile cells but not in other cellular locales (Fig. 2A). Depolymerization of the actin cytoskeleton with the actin inhibitor cytochalasin D (5 μg/mL) removed LKB1 from the leading edge, but LKB1 still seemed associated with a depolymerized actin cytoskeleton (Fig. 2B). Further temporal studies showed that LKB1 associated with actin within 5 min postwounding (Fig. 2C), indicating that this event is rapid and closely tied to its leading edge polarization, which also occurred within 5 min (Fig. 1C).

Lastly, we also examined LKB1 localization in Madin-Darby canine kidney (MDCK) cells, which are noncancerous cell lines that polarize to form an apical and basolateral region in tissue culture dishes when grown to confluency. In nonpolarized cells, endogenous LKB1 is primarily localized in the nucleus and perinuclear region (Supplementary Fig. S1); however, in confluent polarized cells, LKB1 drastically relocalizes to the plasma membrane (Supplementary Fig. S1). Moreover, in the polarized cells, LKB1 colocalizes with actin at the cellular edges (Supplementary Fig. S1), similar to our previous results in NSCLC (Fig. 2). Thus, like the NSCLC cells tested above, LKB1 relocalizes near the plasma membrane and associates with actin when the cell polarity program is triggered in MDCK cells.

LKB1 loss leads to defective Golgi positioning, inhibition of lamellipodia formation, and aberrant lung cancer polarity. Next, we asked the question, does LKB1 regulate NSCLC polarity?
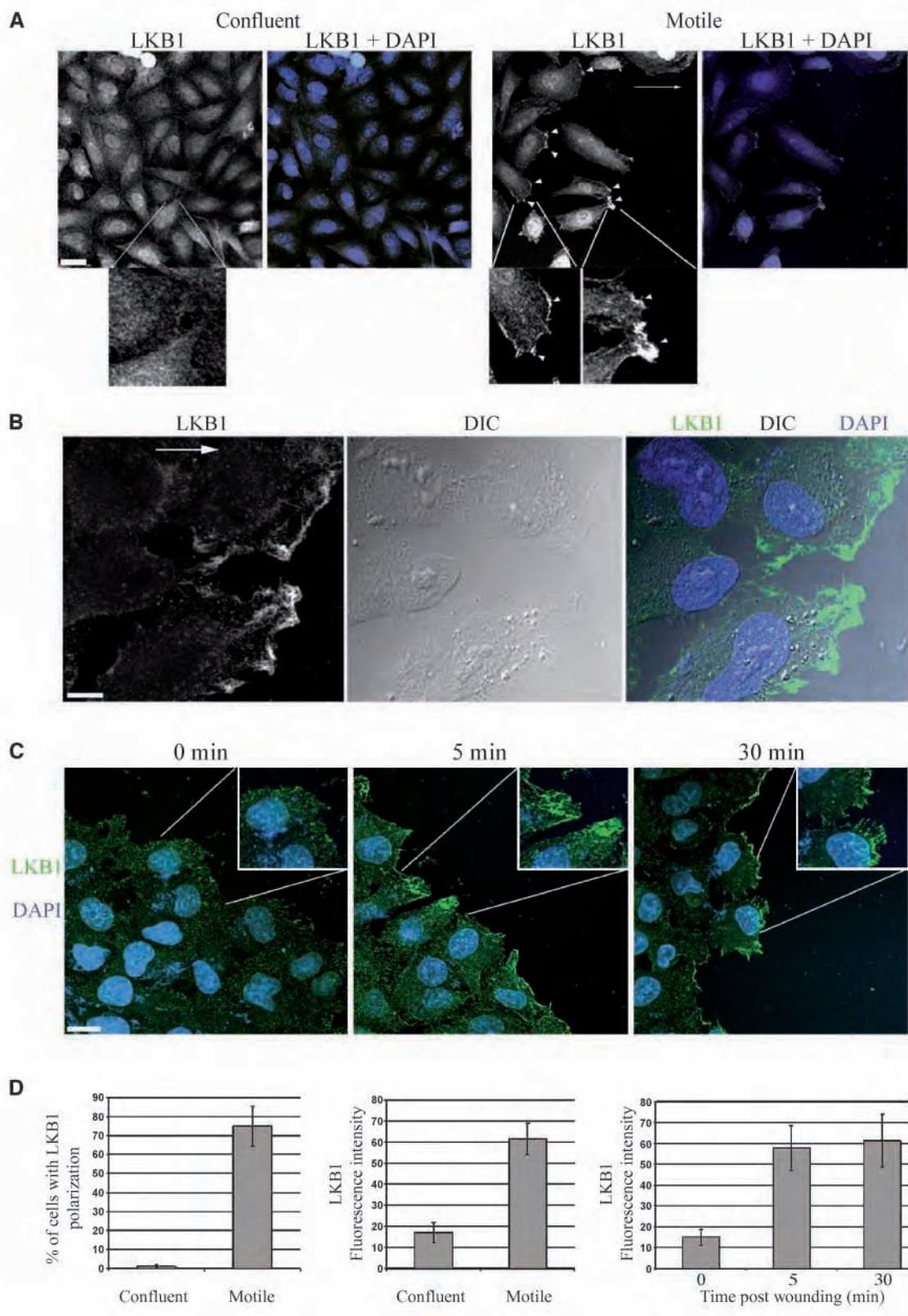
Figure 1. Endogenous LKB1 rapidly polarizes to the leading edge of motile NSCLC cells (*A–C*). Confocal images of endogenous LKB1 immunofluorescence in H1703 NSCLC cells. *Arrow,* direction of movement due to wounding. *A,* LKB1 and DAPI (nuclear) staining in confluent and motile cells. *Arrowheads,* regions of LKB1 polarization. *Bar,* 20 μm. *B,* higher magnification of LKB1 polarization and its relationship to the plasma membrane as revealed by DIC staining. *Blue,* DAPI staining. *Scale bar,* 5 μm. *C,* time course of LKB1 polarization. *Scale bar,* 10 μm. *D,* left bar graph shows the percentage of cells with LKB1 polarization after wounding. Center and right bar graphs show LKB1 mean fluorescent intensity; *n* = 20 cells per experimental group; *bars,* SD.
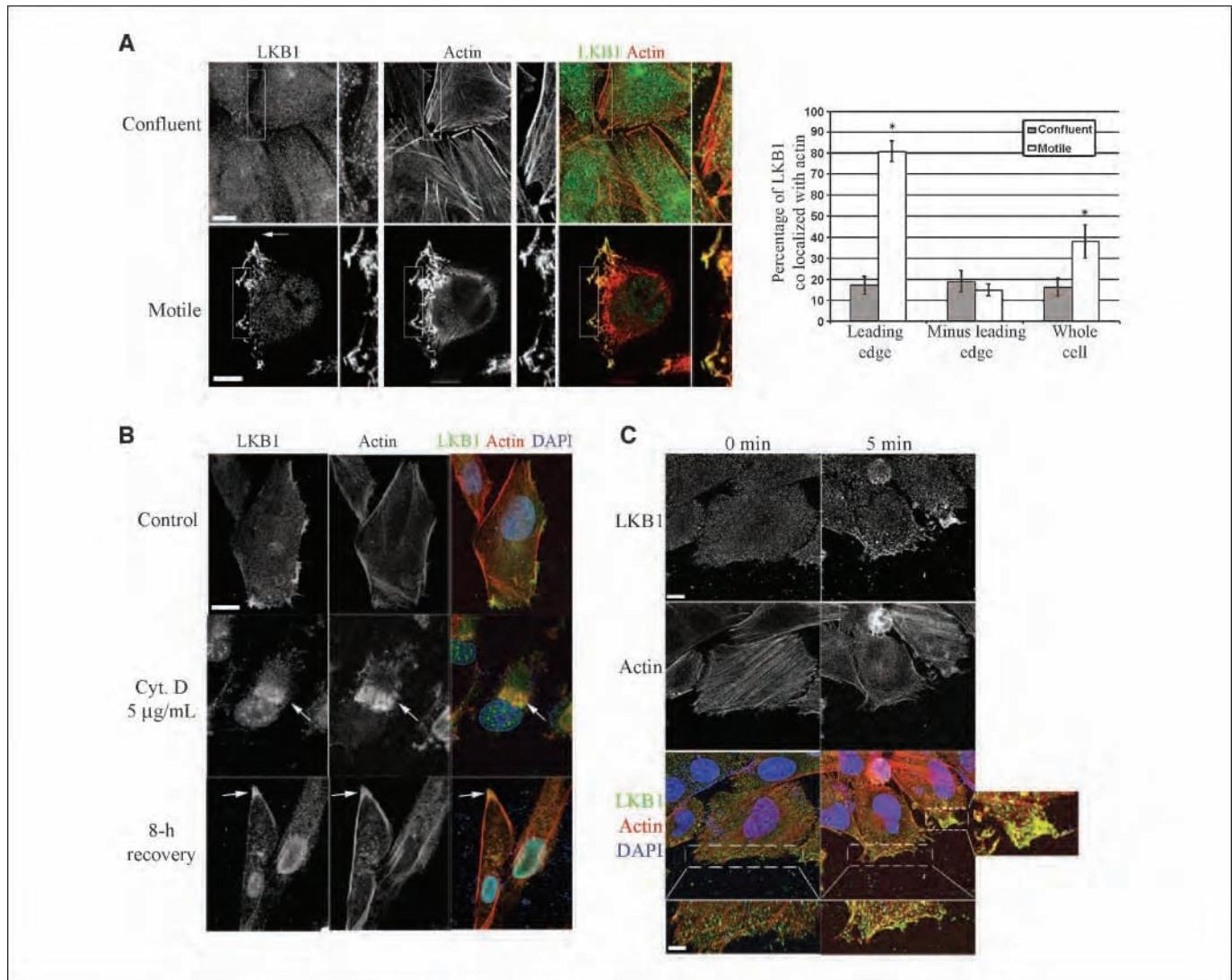
Figure 2. LKB1 associates with actin in motile NSCLC cells (*A–C*). Confocal images of LKB1, actin, and DAPI immunofluorescence. *A,* LKB1 and actin staining in confluent and motile cells. Insets show magnified view of boxed region. Bar graph shows percentage colocalization of LKB1 and actin in the various cellular regions (*n* = 10 cells). *Scale bar,* 10 μm. *Bars,* SD. *B,* LKB and actin staining in cytochalasin D–treated cells and after an 8-h recovery from treatment. *Scale bar,* 10 μm. *C,* LKB1 and actin staining at 0 and 5 min postwounding. *Scale bars,* 5 μm. Insets show magnified field of boxed region; *scale bar,* 2 μm.

To do this, we used siRNA to deplete LKB1 and subsequently examined the effect on cell polarity. Due to the Golgi's role in protein trafficking to the leading edge, it serves as a marker of cell polarization and reorients between the nucleus and the leading edge when cell polarity is stimulated (18, 19). We therefore examined Golgi reorientation in a quantifiable manner by dividing cells into three regions (120° in each region), such that in unpolarized cells the Golgi had an equal likelihood of being in any of the three regions (example in Supplementary Fig. S2), but in polarized cells, the Golgi realigns to the 120° region facing the direction of movement (18, 19). To assess this, staining of the Golgi, LKB1, and nucleus was performed in control siRNA–transfected and LKB1 siRNA–transfected cells (Fig. 3*A*). In cells transfected with control siRNA, the Golgi realigned by 6 h, with nearly 70% of cells having proper Golgi orientation. In contrast, LKB1 depleted cells showed only 35% of cells with proper Golgi positioning after 6 h and only 45% after 16 h, indicating defective cell polarity in the absence of LKB1 (Fig. 3*B* and *C*). Additional defects in cell polarity were also

observed when cell shape and lamellipodia formation were analyzed. Control cells were elongated perpendicular to the wound front and had a clear lamellipodia by 6 h, whereas LKB1-depleted cells remained randomly oriented and lacked a distinctive lamellipodia (Fig. 3*B*). Image quantitation (example in Supplementary Fig. S2) showed that controls cells repositioned nearly 70° relative to the wound by 6 h and maintained this position for at least 16 h (Fig. 3*D, top and bottom*). In contrast, LKB1-depleted cells had a 30° orientation relative to the wound at 6 h and 15° orientation by 16 h. Interestingly, LKB1-depleted cells ultimately aligned parallel to the wound instead of the normal perpendicular orientation observed in control cells (Supplementary Fig. S2), further indicating defective cell polarity. Similar results were also obtained in other NSCLC cell lines with wild-type LKB1 (data not shown), indicating that this is not a cell line–specific event. Taken together, these data indicate that LKB1 is necessary for NSCLC polarity.

LKB1 colocalizes with the polarity proteins cdc42 and PAK and regulates cdc42 recruitment. Cdc42 is a master regulator of

cell polarity and also polarizes to the leading edge of motile cells (2); therefore, we next determined the relationship between LKB1 and cdc42 in this region. Immunofluorescence colocalization analysis of endogenous LKB1 and cdc42 in H1299 cells revealed that a significant portion of LKB1 colocalized with cdc42 within 15 min postwounding at the leading edge (Fig. 4A). Interestingly, the two proteins were not associated in any other regions within the cytoplasm (Fig. 4A) and LKB1 only colocalized with cdc42 in motile but not confluent cells (data not shown). Similar colocalization between LKB1 and cdc42 were also observed in H1703 NSCLC cells (Supplementary Fig. S3).

We next examined LKB1 colocalization with the cdc42 binding partner PAK. When active, cdc42 binds to PAK causing PAK phosphorylation and eliciting a variety of downstream cell polarity events, including Golgi reorientation and actin remodeling (4, 20). Colocalization analysis showed that, indeed, LKB1 colocalizes with PAK only at the leading edge (Fig. 4A), suggesting that LKB1, cdc42, and PAK are possibly linked in the cell polarity pathway. Moreover, LKB1 also colocalizes at the leading edge with phosphorylated PAK (Ser[144]; Supplementary Fig. S3), which represents the active form of

PAK. To exclude the possibility of a putative cell adhesion role for LKB1, we stained for phosphorylated focal adhesion kinase (FAK), a functional marker of the cell adhesion pathway. In this case, LKB1 did not colocalize phosphorylated FAK, rather it was excluded from these adhesive sites (Supplementary Fig. S3), suggesting that LKB1 is not involved in the cell adhesion pathway or localizes to adhesive sites within the motile cell.

To further examine the spatiotemporal relationship between LKB1 and cdc42, we performed a 10-min short-time course of LKB1 and cdc42 localization to determine when these proteins colocalize. These results showed that LKB1 accumulates near the plasma membrane as early as 1 min postwounding, and this accumulation precedes cdc42 arrival (Fig. 4B). By 5 min, cdc42 begins to accumulate at the leading edge, at LKB1-marked sites (Fig. 4B). This trend continues, and by 10 min, both proteins are colocalized (Fig. 4B). Thus, LKB1 polarization precedes that of cdc42, but ultimately cdc42 colocalizes to the LKB1-designated sites. Notably, we did not observe any cells where cdc42 was localized to the leading edge without LKB1 already there; however, there were cells that had neither LKB1 nor cdc42 at the leading edge at the early
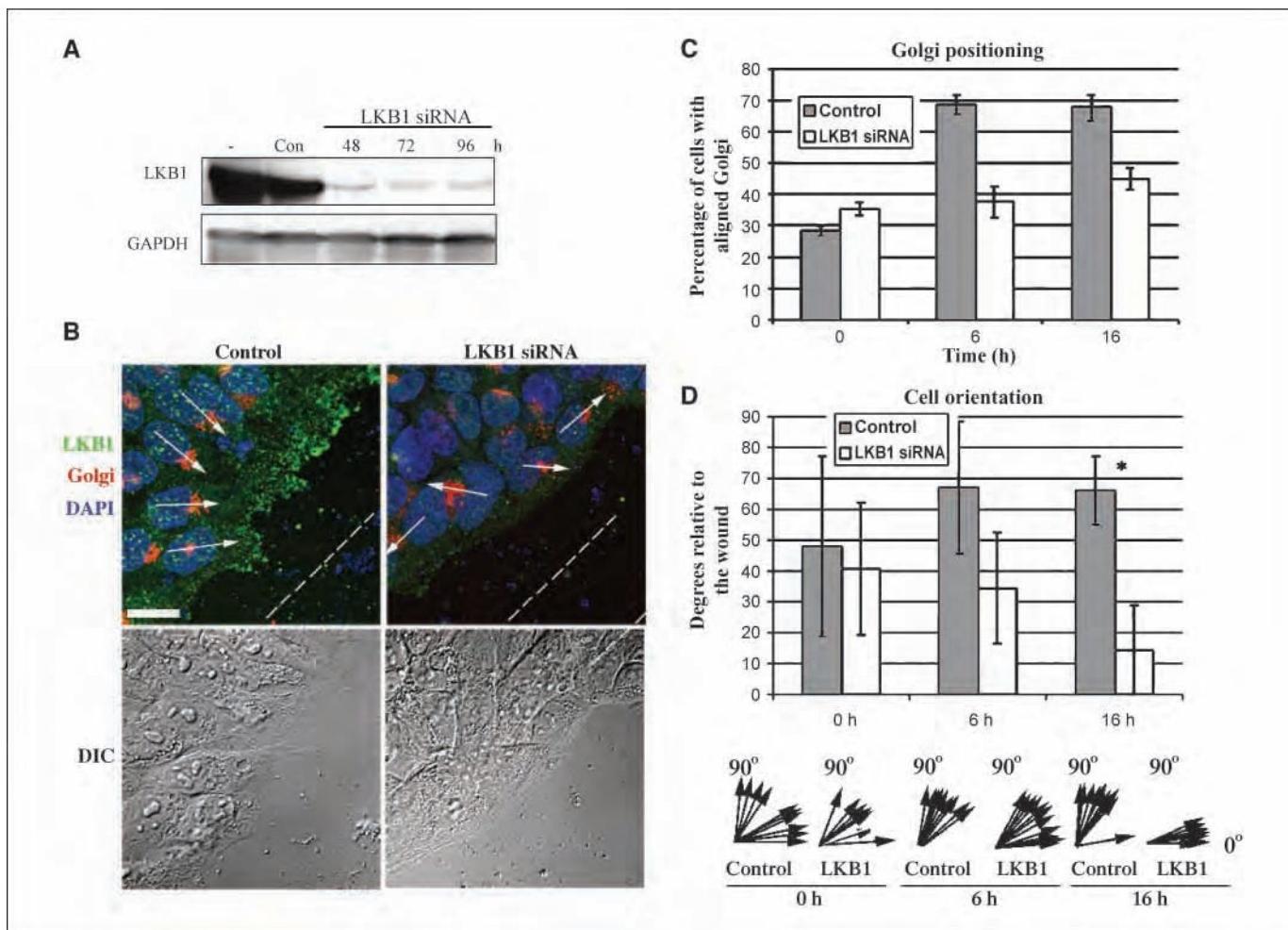
Figure 3. LKB1 depletion causes defective cell polarity. *A,* Western blot showing siRNA depletion of LKB1 in H1792 NSCLC cells. *Dash,* no transfection; *Con,* control siRNA. *B,* LKB1, Golgi, and DAPI immunofluorescence in H1792 NSCLC cells. *Arrows,* Golgi positioning relative to the wound designated with a dashed line. DIC images below show cellular morphology and lamellipodia of the same cells as above. *Scale bar,* 10 μm. *C,* bar graph quantitating the percentage of cells with aligned Golgi in control and LKB1-depleted cells (*n* = 20 cells per experimental group). *D, top,* bar graph depicting cell positioning in degrees relative to the wound (*n* = 20 cells per experimental group; *bottom*). *Arrows,* graphical representation of cellular positioning relative to the wound in control and LKB1-depleted cells.
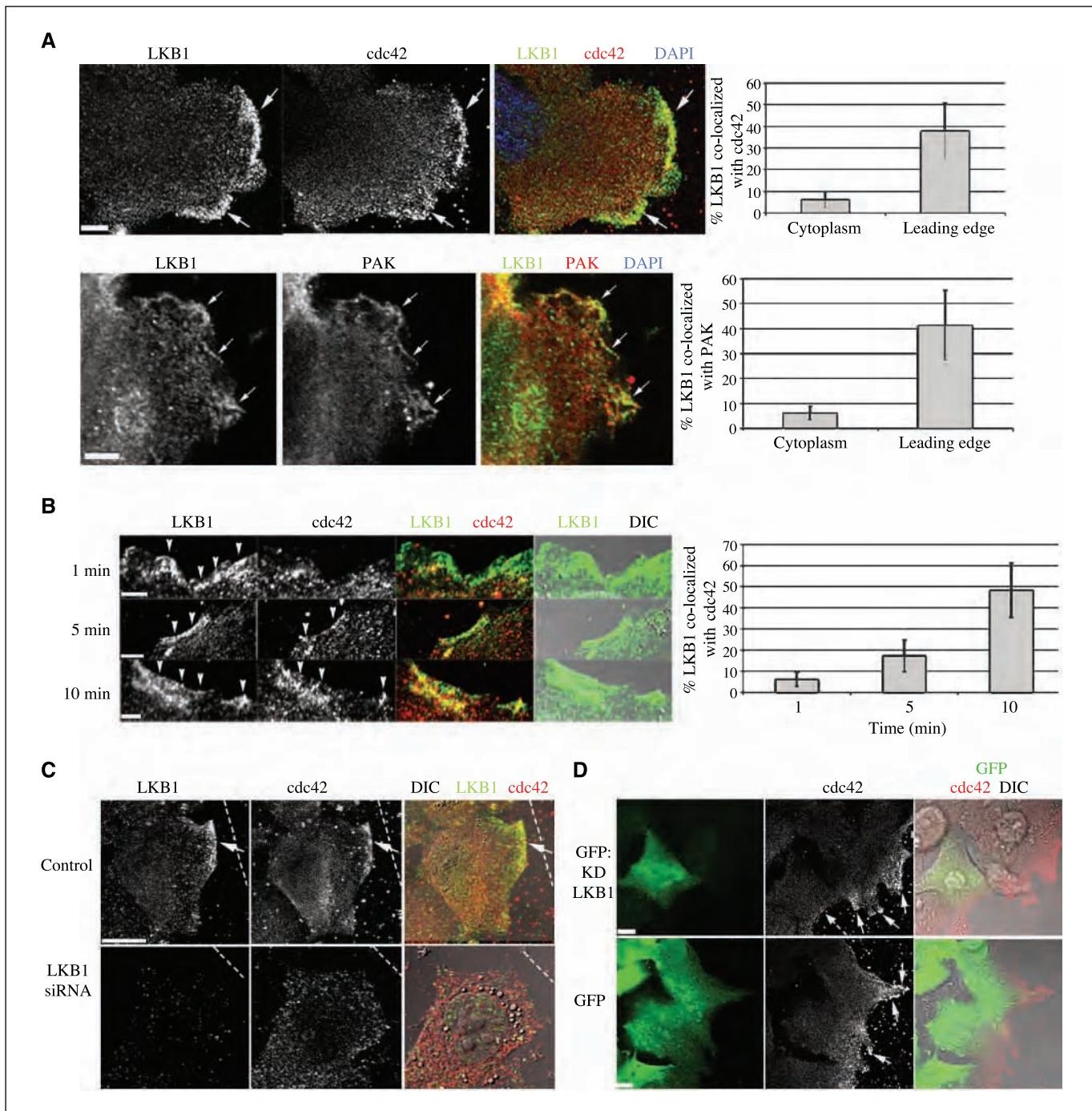
Figure 4. LKB1 colocalizes with cdc42, PAK, and p-PAK and precedes cdc42 recruitment (*A–D*). Confocal images of LKB and cdc42 or PAK immunofluorescence in H1299 NSCLC cells. Bar graphs show percentage LKB1 colocalized with the respective protein. In all cases, error bars are SD. *A,* LKB1 and cdc42 staining at the leading edge (*top*). *Arrows,* regions of colocalization. LKB1 and PAK staining at the leading edge (*bottom*). *Arrows,* regions of colocalization. *Scale bars,* 5 μm. *B,* short-interval time course of LKB1 and cdc42 localization. DIC staining shows LKB1 localization at or near the plasma membrane. *Scale bar,* 2 μm. *C,* LKB1 and cdc42 staining in control and LKB1 siRNA knockdown cells. *Arrows,* cdc42 polarization to the leading edge; *dashed lines,* the wound front. *Scale,* 10 μm. *D,* confocal images of cdc42 staining in a cell expressing either a dominant-negative kinase dead LKB1 (GFP:KD LKB1) or GFP control (*bottom*). *Arrow,* cdc42 polarization; *dashed arrow,* lack of cdc42 polarization. Insets show magnified view. *Scale,* 10 μm. For all bar graphs, *n* = 10 cells per experimental group.

time points. This same pattern of LKB1-cdc42 colocalization was observed in H1703 NSCLC cells (Supplementary Fig. S3). Taken together, these results suggest that LKB1 serves as an intracellular marker for future cdc42 recruitment.

The above results suggest that LKB1 may regulate cdc42 recruitment. To test this, LKB1 was depleted in H1299 cells with

siRNA and cdc42 recruitment to the leading edge was assayed. In cells transfected with control siRNA, LKB1 and cdc42 were colocalized at the leading edge after 15 min postwounding in nearly all cells observed as expected (Fig. 4*C*). In contrast, in cells transfected with LKB1 siRNA, cdc42 did not polarize to the leading edge but rather displayed a punctate appearance throughout the

cytoplasm (Fig. 4*C*). As an alternative approach, we overexpressed a kinase dead mutant of LKB1 fused to GFP (GFP:KD LKB1; ref. 21) in LKB1 wild-type cells. This GFP:KD LKB1 lacks LKB1 kinase activity and therefore serves as an excellent probe for understanding the specific role of LKB1 kinase activity. In GFP:KD LKB1–positive cells, cdc42 did not polarize to the leading edge (Fig. 4*D*), whereas neighboring cells that were not expressing the kinase dead LKB1 did, indeed, have cdc42 polarization (Fig. 4*D*). Furthermore, these cells did not have a lamellipodia, consistent with the idea that LKB1 kinase activity is required for its role in cell polarity. Transfections with a control GFP-only vector did not alter cdc42 recruitment or lamellipodia formation (Fig. 4*D*). Thus, when LKB1 expression is depleted or endogenous LKB1 activity is inhibited with a kinase dead dominant-negative LKB1, cdc42 cannot polarize to the leading edge. These results therefore support the hypothesis that LKB1 is required for cdc42 recruitment to the leading edge.

LKB1 mediates cdc42 activity by complexing with PAK and active cdc42, but not inactive cdc42. Previous reports show that cdc42 is recruited and activated at the leading edge (22, 23) by a series of GTPase-activating proteins (GAP) and guanine exchange factors (GEF; ref. 22). Because we show that LKB1 is essential for cdc42 leading edge recruitment, we assayed for active cdc42 levels in control siRNA–transfected and LKB1 siRNA–transfected cells prewounding and postwounding, using a coimmunoprecipitation approach with PAK-coated beads. This immunoprecipitation-based approach is based upon the principle that active cdc42 (GTP bound) binds specifically to the PBD domain of PAK. Thus, this assay uses PAK-PBD agarose beads to selectively isolate and pull down active cdc42 (cdc42-GTP) from lysates. Subsequently, the precipitated GTP-cdc42 is detected by Western blot analysis using a mouse anti-cdc42 antibody. To activate cdc42, multiple wounds were made to confluent H1299 NSCLC cells on 10 cm² dishes; this is the standard approach for cdc42 activation (Fig. 5*A*; ref. 19). Control H1299 LKB1 wild-type NSCLC cells initially have low prewounding levels of active cdc42, but after 10 min postwounding, active cdc42 substantially increases as predicted (Fig. 5*B*).

In contrast, in LKB1-depleted cells, active cdc42 levels were significantly lower, in both prewounding and postwounding samples relative to their control counterparts (Fig. 5*B*). However, cdc42 activation itself was not impaired, because we observed a nearly 2-fold increase in cdc42 levels in control and LKB1-depleted cells (Fig. 5*B*). These data show that LKB1 depletion significantly reduces active cdc42 levels, but not cdc42 activation per se. Total cdc42 levels did not change throughout the experiment (Fig. 5*B*). Similar experiments were performed with the kinase dead GFP:KD LKB1 and showed that the levels of active cdc42 postwounding were significantly reduced in the population expressing the kinase dead LKB1 compared with the wild-type LKB1 (Fig. 5*C*). Thus, this result is consistent with LKB1 mediating active cdc42 levels.

Because these results suggest that LKB1 can mediate cdc42 activity levels, we wanted to determine if LKB1 complexes with cdc42 and the downstream cdc42 binding partner PAK. Moreover, we also wanted to examine the possibility that a putative association is dependent upon the activity state of cdc42. Therefore, we performed coimmunoprecipitation assays in H1299 cells with co-transfected FLAG-LKB, myc-PAK, and three different GFP-cdc42 constructs — a constitutively active cdc42 (Cdc42-Q61L), dominant-negative cdc42 (Cdc42-T17N), or wild-type cdc42 (23, 24). The activity of the Cdc42-Q61L and inactivity of cdc42-T17N was validated in our wounding assay (Supplementary Fig. S4). Importantly, these results show that LKB1 has an enhanced association with the constitutively active cdc42-Q61L mutant compared with inactive cdc42-T17N or wild-type cdc42 (presumably a mixed population of active and inactive cdc42; Fig. 5*C*). Specifically, the cdc42-Q61L mutant coimmunoprecipitated with LKB1, whereas the cdc42-T17N and wild-type cdc42 showed minimal association. In addition, myc-PAK showed an enhanced coimmunoprecipitation with LKB1 in the presence of the constitutively active cdc42-Q61L mutant, suggesting that LKB1, PAK, and active cdc42 are complexed together within the cell in a cdc42 activity–dependent manner (Fig. 5*D*).
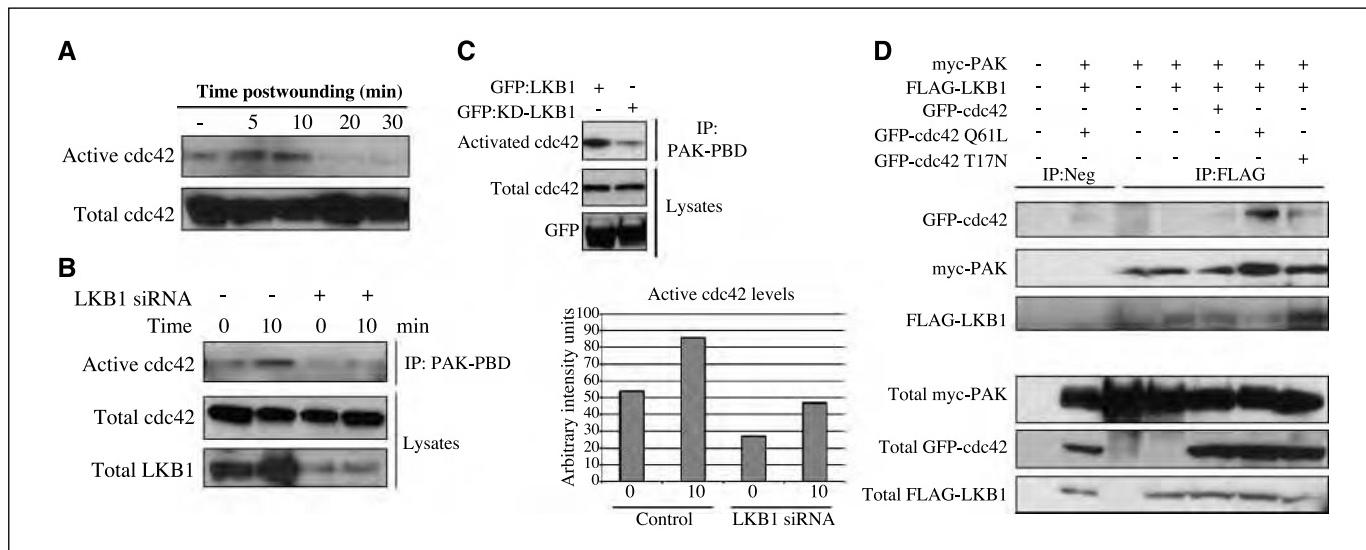
Figure 5. LKB1 mediates active cdc42 levels and associates with the active cdc42-PAK complex. *A,* Western blot showing a time course of cdc42 activation in our wounding model. *B,* Western blot showing cdc42 activation in control and LKB1-depleted cells at 0 min (prewounding) and 10 min (postwounding). *C,* Western blot of a cdc42 activation assay in H1299 NSCLC cells overexpressing a dominant-negative GFP:KD LKB1 after 15 min postwounding compared with wild-type GFP:LKB1. *D,* immunoblotting of FLAG-LKB1 immunoprecipitation in H1299 cells transfected with the various constructs shown. Immunoblots of total lysates are shown below.
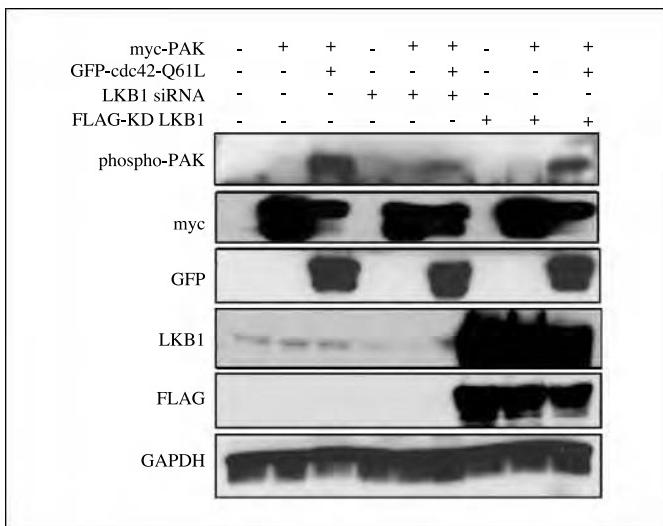
Figure 6. Compromised LKB1 functionality affects downstream PAK phosphorylation. Western blot of phosphorylated PAK levels in control, LKB1-depleted cells, and cells transfected with a dominant-negative FLAG-KD LKB1. A GFP-labeled constitutively active cdc42 mutant (GFP-cdc42-Q61L) was overexpressed to activate myc-PAK phosphorylation.

Compromised LKB1 functionality affects downstream PAK phosphorylation. Binding of active cdc42 to PAK stimulates its phosphorylation and triggers various cell polarity event (4). To determine the effect of LKB1 on the downstream activation of PAK, levels of phosphorylated PAK in control and LKB1-depleted cells were examined. In control cells, phosphorylated PAK levels are initially low, but are stimulated in the presence of constitutively active cdc42-Q61L mutant as expected (Fig. 6A). In contrast, the cdc42-Q61L–induced stimulation of PAK phosphorylation was suppressed in LKB1-depleted cells or cells overexpressing a kinase dead LKB1 mutant (GFP:KD LKB1; Fig. 6A). Taken together, these results indicate that LKB1 activity is required for active cdc42 to induce PAK phosphorylation. Furthermore, LKB1 also colocalizes with phosphorylated PAK sites at the leading edge but not throughout the cytoplasm (Supplementary Fig. S3). This is similar to LKB1 localization with both cdc42 and total PAK (Supplementary Fig. S3); however, phosphorylated PAK seems to form well-defined puncta within the LKB1-stained regions.

Discussion

LKB1 behaves as a dynamic protein that rapidly migrates to the cellular leading edge within 5 min and associates with actin. This pattern of localization is similar to other cell polarity proteins, such as the small Rho GTPase rac, β-catenin, and GSK3β (17, 25, 26). Moreover, a polarized LKB1 localization was recently observed in some human lung cancer tissues by immunohistochemistry (27). LKB1 polarization and activity seem to be regulated by the cell polarity stimulus because LKB1 has a diffused cytoplasmic and nuclear pattern before wounding, but upon cell polarity stimulation, LKB1 relocalizes to the leading edge and colocalizes with actin filaments. LKB1 rarely colocalizes with actin stress fibers that span the cytoplasm, suggesting that the mechanism of LKB1 translocation to the leading edge is not actin-based; however, it requires actin to maintain its localization to the leading edge because actin inhibition removes LKB1 from this region.

LKB1 is essential for NSCLC polarity in our wounding model because LKB1 depletion results in classic cell polarity defects, such as aberrant Golgi positioning, reduced lamellipodia formation, and aberrant morphology. Previous studies show that these cell polarity events are regulated, at least in part, by the small rho GTPase cdc42 and its downstream binding partner PAK (2). Importantly, our data now show that LKB1 is a critical member of this cdc42-PAK pathway. LKB1 protein expression and kinase activity are essential for both the recruitment and maintenance of active cdc42 levels. Specifically, when LKB1 function is compromised, cdc42 does not polarize to the leading edge and active cdc42 levels are significantly diminished. Cdc42 activation involves a series of GAPs and GEFs (22) As cdc42 activation takes place at the leading edge (23), any defect in its polarization to the leading edge would be expected to affect levels of cdc42 activation. Our data are most consistent with the idea that LKB1 is necessary for the translocation of cdc42 rather than its direct activation as wounding induces a similar 2-fold activation of cdc42 in both control and LKB1-depleted cells.

LKB1 forms a complex with cdc42 and PAK in NSCLC, and this interaction is dependent on the activation status of cdc42. Consistent with this idea, LKB1 colocalizes with cdc42, PAK, and phosphorylated PAK at the cellular leading edge when the cell polarity program is stimulated but does not colocalize with any of these proteins before cell polarity stimulation or in other regions throughout the cytoplasm. Furthermore, LKB1 depletion or inhibition of its kinase activity suppresses PAK phosphorylation at Ser[144], a site critical for activation of PAK kinase activity (4). Active cdc42 is known to stimulate PAK phosphorylation, leading to downstream induction of cell morphogenesis, changes in actin/microtubule dynamics, and cell motility (4). Thus, the cell polarity defects observed in the absence of LKB1 are thus likely due to impaired translocation and activation of cdc42 and consequently diminished PAK phosphorylation. Interestingly, initial results show that cell lines naturally defective in LKB1 (e.g., A549) have reduced cdc42-dependent PAK phosphorylation (data not shown), suggesting that natural loss of LKB1 function impairs PAK phosphorylation.

Taken together, we propose an LKB1-cdc42-PAK pathway such that, when cell motility and polarity are stimulated, LKB1 polarizes to the leading edge where it associates with actin to serve as a scaffold for the subsequent recruitment and activation of cdc42. Once this occurs, a complex between LKB1, active cdc42, and PAK assembles, resulting in PAK phosphorylation and downstream activation of cell polarity events, such as lamellipodia formation and Golgi reorientation. We believe that LKB1 is an essential component of this process, because LKB1 depletion leads to reduced cdc42 activation and recruitment, reduced PAK phosphorylation, and negatively affects several hallmarks of normal cell polarity.

It is interesting to note that PAK is a member of the STE20 family of kinases (28), as well as the LKB1 cofactor binding partner STRAD, although the kinase domain of STRAD lacks several residues indispensable for intrinsic catalytic activity and is, thus, catalytically inactive (29). Because the interaction between STRAD and LKB1 is necessary for LKB1 kinase activity and localization (30), this structural relatedness between PAK and STRAD may provide additional interesting clues on the functional consequences of the LKB1-PAK association. Lastly, LKB1 is also is an upstream kinase to AMP-activated kinase (AMPK), which serves as a sensor for energy stress (31). Recent studies show that AMPK also functions in cell polarity (32–34) and AMPK mutants lose their polarity and overproliferate under energetic stress (32). Thus, AMPK may

indeed be a downstream substrate of the LKB1-cdc42-PAK pathway, and future studies will investigate this.

More than 30% of NSCLC tumors harbor mutations in LKB1 (11, 12). This high LKB1 mutation rate, in combination with the data presented herein, provides the intriguing possibility that LKB1 loss in human tumors could cause aberrant lung cancer cell polarity. In fact, C-terminal LKB1 mutations affect the ability of intestinal epithelial cells and migrating astrocytes to establish and maintain polarity (14). Because a loss of epithelial cell polarity is linked to increased cancer invasion via EMT (6, 35), LKB1 mutations in patients may serve as a trigger for cancer cell invasion. This idea is supported by recent work, showing that a somatically activatable mutant Kras-driven model of mouse lung cancer showed that homozygous inactivation of LKB1 resulted in shorter latency and more frequent metastasis compared with tumors lacking p53 or Ink4a/Arf (36).

Moreover, cells expressing mutant LKB1 possessed greater invasive potential compared with wild-type LKB1 cells (37). Thus, future experiments will address this possibility and determine if a compromised LKB1-cdc42-PAK pathway triggers cancer cell invasion.

Acknowledgments

References

1. Vicente-Manzanares M, Webb DJ, Horwitz AR. Cell migration at a glance. J Cell Sci 2005;118:4917–9.
2. Etienne-Manneville S. Cdc42-the centre of polarity. J Cell Sci 2004;117:1291–300.
3. Manser E, Leung T, Salihuddin H, Zhao ZS, Lim L. A brain serine/threonine protein kinase activated by Cdc42 and Rac1. Nature 1994;367:40–6.
4. Bokoch GM. Biology of the p21-activated kinases. Annu Rev Biochem 2003;72:743–81.
5. Lei M, Lu W, Meng W, et al. Structure of PAK1 in an autoinhibited conformation reveals a multistage activation switch. Cell 2000;102:387–97.
6. Thiery JP. Epithelial-mesenchymal transitions in development and pathologies. Curr Opin Cell Biol 2003;15:740–6.
7. Jenne DE, Reimann H, Nezu J, et al. Peutz-Jeghers syndrome is caused by mutations in a novel serine threonine kinase. Nat Genet 1998;18:38–43.
8. Alessi DR, Sakamoto K, Bayascas JR. LKB1-dependent Signaling Pathways. Annu Rev Biochem 2006;75:137–63.
9. Shaw RJ, Bardeesy N, Manning BD, et al. The LKB1 tumor suppressor negatively regulates mTOR signaling. Cancer Cell 2004;6:91–9.
10. Baas AF, Smit L, Clevers H. LKB1 tumor suppressor protein: PARtaker in cell polarity. Trends Cell Biol 2004;14:312–9.
11. Carretero J, Medina PP, Pio R, Montuenga LM, Sanchez-Cespedes M. Novel and natural knockout lung cancer cell lines for the LKB1/STK11 tumor suppressor gene. Oncogene 2004;23:4037–40.
12. Sanchez-Cespedes M, Parrella P, Esteller M, et al. Inactivation of LKB1/STK11 is a common event in adenocarcinomas of the lung. Cancer Res 2002;62:3659–62.
13. Baas AF, Kuipers J, van der Wel NN, et al. Complete polarization of single intestinal epithelial cells upon activation of LKB1 by STRAD. Cell 2004;116:457–66.
14. Forcet C, Etienne-Manneville S, Gaude H, et al. Functional analysis of Peutz-Jeghers mutations reveals that the LKB1 C-terminal region exerts a crucial role in regulating both the AMPK pathway and the cell polarity. Hum Mol Genet 2005;14:1283–92.
15. Schafer-Hales K, Iaconelli J, Snyder JP, et al. Farnesyl transferase inhibitors impair chromosomal maintenance in cell lines and human tumors by compromising CENP-E and CENP-F function. Mol Cancer Ther 2007;6:1317–28.
16. Marcus AI, Zhou J, O'Brate A, et al. The synergistic combination of the farnesyl transferase inhibitor lonafarnib and paclitaxel enhances tubulin acetylation and requires a functional tubulin deacetylase. Cancer Res 2005;65:3883–93.
17. Etienne-Manneville S, Hall A. Cdc42 regulates GSK-3β and adenomatous polyposis coli to control cell polarity. Nature 2003;421:753–6.
18. Chacko AD, Hyland PL, McDade SS, Hamilton PW, Russell SH, Hall PA. SEPT9_v4 expression induces morphological change, increased motility and disturbed polarity. J Pathol 2005;206:458–65.
19. Etienne-Manneville S, Hall A. Integrin-mediated activation of Cdc42 controls cell polarity in migrating astrocytes through PKCζ. Cell 2001;106:489–98.
20. Raftopoulou M, Hall A. Cell migration: Rho GTPases lead the way. Dev Biol 2004;265:23–32.
21. Karuman P, Gozani O, Odze RD, et al. The Peutz-Jegher gene product LKB1 is a mediator of p53-dependent cell death. Mol Cell 2001;7:1307–19.
22. Itoh RE, Kurokawa K, Ohba Y, Yoshizaki H, Mochizuki N, Matsuda M. Activation of rac and cdc42 video imaged by fluorescent resonance energy transfer-based single-molecule probes in the membrane of living cells. Mol Cell Biol 2002;22:6582–91.
23. Nalbant P, Hodgson L, Kraynov V, Toutchkine A, Hahn KM. Activation of endogenous Cdc42 visualized in living cells. Science 2004;305:1615–9.
24. Subauste MC, Von Herrath M, Benard V, et al. Rho family proteins modulate rapid apoptosis induced by cytotoxic T lymphocytes and Fas. J Biol Chem 2000;275:9725–33.
25. Nobes CD, Hall A. Rho GTPases control polarity, protrusion, and adhesion during cell movement. J Cell Biol 1999;144:1235–44.
26. Kraynov VS, Chamberlain C, Bokoch GM, Schwartz MA, Slabaugh S, Hahn KM. Localized Rac activation dynamics visualized in living cells. Science 2000;290:333–7.
27. Conde E, Suarez-Gauthier A, Garcia-Garcia E, et al. Specific pattern of LKB1 and phospho-acetyl-CoA carboxylase protein immunostaining in human normal tissues and lung carcinomas. Hum Pathol 2007;38:1351–60.
28. Dan I, Watanabe NM, Kusumi A. The Ste20 group kinases as regulators of MAP kinase cascades. Trends Cell Biol 2001;11:220–30.
29. Baas AF, Boudeau J, Sapkota GP, et al. Activation of the tumour suppressor kinase LKB1 by the STE20-like pseudokinase STRAD. EMBO J 2003;22:3062–72.
30. Boudeau J, Baas AF, Deak M, et al. MO25α/β interact with STRADα/β enhancing their ability to bind, activate and localize LKB1 in the cytoplasm. EMBO J 2003;22:5102–14.
31. Hawley SA, Boudeau J, Reid JL, et al. Complexes between the LKB1 tumor suppressor, STRAD α/β and MO25 α/β are upstream kinases in the AMP-activated protein kinase cascade. J Biol 2003;2:28.
32. Mirouse V, Swick LL, Kazgan N, St Johnston D, Brenman JE. LKB1 and AMPK maintain epithelial cell polarity under energetic stress. J Cell Biol 2007;177:387–92.
33. Zhang L, Li J, Young LH, Caplan MJ. AMP-activated protein kinase regulates the assembly of epithelial tight junctions. Proc Natl Acad Sci U S A 2006;103:17272–7.
34. Zheng B, Cantley LC. Regulation of epithelial tight junction assembly and disassembly by AMP-activated protein kinase. Proc Natl Acad Sci U S A 2007;104:819–22.
35. Huber MA, Kraut N, Beug H. Molecular requirements for epithelial-mesenchymal transition during tumor progression. Curr Opin Cell Biol 2005;17:548–58.
36. Ji H, Ramsey MR, Hayes DN, et al. LKB1 modulates lung cancer differentiation and metastasis. Nature 2007;448:807–10.
37. Upadhyay S, Liu C, Chatterjee A, et al. LKB1/STK11 suppresses cyclooxygenase-2 induction and cellular invasion through PEA3 in lung cancer. Cancer Res 2006;66:7870–9.

From combinatorial peptide selection to drug prototype (I): Targeting the vascular endothelial growth factor receptor pathway

Ricardo J. Giordano[a,1,2], Marina Cardó-Vila[a,1], Ahmad Salameh[a], Cristiane D. Anobom[b], Benjamin D. Zeitlin[c], David H. Hawke[d], Ana P. Valente[b], Fábio C. L. Almeida[b], Jacques E. Nör[c], Richard L. Sidman[e,3], Renata Pasqualini[a,3], and Wadih Arap[a,3]

[a]David H. Koch Center, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030; [b]National Nuclear Magnetic Resonance Center, Rio de Janeiro Federal University, Rio de Janeiro 21941, Brazil; [c]Angiogenesis Research Laboratory, University of Michigan School of Dentistry, Ann Arbor, MI 48109; [d]Center for Targeted Therapy, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030; and [e]Harvard Medical School and Department of Neurology, Beth Israel Deaconess Medical Center, Boston, MA 02215

Inhibition of blood vessel formation is a viable therapeutic approach in angiogenesis-dependent diseases. We previously used a combinatorial screening on vascular endothelial growth factor (VEGF)-activated endothelial cells to select the sequence CPQPRPLC and showed that the motif Arg-Pro-Leu targets VEGF receptor-1 and neuropilin-1. Here, we evaluated and validated $_D$(LPR), a derivative molecule with strong antiangiogenesis attributes. This prototype drug markedly inhibits neovascularization in three mouse models: Matrigel-based assay, functional human/murine blood vessel formation, and retinopathy of prematurity. In addition to its systemic activity, $_D$(LPR) also inhibits retinal angiogenesis when administered in an eye-drop formulation. Finally, in preliminary studies, we have showed targeted drug activity in an experimental tumor-bearing mouse model. These results show that drugs targeting extracellular domains of VEGF receptors are active, affect signal transduction, and have potential for clinical application. On a larger context, this study illustrates the power of ligand-directed selection plus retro-inversion for rapid drug discovery and development.

peptide | cancer | VEGFR | angiogenesis | retinopathy of prematurity

The formation of blood vessels, whether de novo (vasculogenesis) or from existing blood vessels (angiogenesis), is a fundamental biological process. In the 1970s, Folkman (1) introduced the concept of angiogenesis-dependent diseases and suggested that compounds inhibiting neovascularization would find applications in medicine, particularly against cancer (1). Although this idea was entertained with skepticism at the time, several inhibitors are currently in clinical use, and others are at advanced stages of development. As such, angiogenesis stands as "an organizing principle for drug discovery" (2) and has led to converging insights into the mechanisms of pathological disorders otherwise considered dissimilar to each other (e.g., tumor growth and diabetes to rheumatoid arthritis and thalidomide-based malformations). We have previously proposed abnormal vascularization to be the one unifying feature of seemingly disparate ocular diseases such as diabetic retinopathy, age-related macular degeneration, and retinopathy of prematurity (3); indeed, abnormal blood-vessel formation in the human retina is a leading cause of blindness from children to adults to the elderly.

A balance of activators and inhibitors coordinates the angiogenic process in health and disease (2, 4, 5). Among the molecular activators, vascular endothelial growth factor (VEGF) and its receptors are considered essential contributors to angiogenesis (6), and both ligands and receptors have been targets of therapies. Since the introduction of the monoclonal antibody bevacizumab (Avastin), the first Food and Drug Administration (FDA)-approved, VEGF-targeted therapy for use in metastatic colorectal cancer in combination with chemotherapy (7), several drugs targeting VEGF-related pathways have been developed and are

currently in various stages of testing (2, 8). FDA approval of pegaptanib (Macugen), an anti-VEGF pegylated aptamer, and Ranibizumab (Lucentis), a fragment of bevacizumab for the treatment of the wet type of age-related macular degeneration, are examples of anti-VEGF agents with proven efficacy toward an eye disorder with an angiogenic component (9, 10). Although VEGF-targeted therapies have shown relative success and improved cancer survival, there are still unsolved issues. For instance, many patients do not respond to VEGF-targeted therapy, and most patients who do respond initially develop resistance (11). Therefore, a better understanding of the mechanisms of action of these agents is necessary and will improve the efficacy of these therapies (12). Also, development of a new generation of agents targeted against other members of the VEGF family might overcome some of the difficulties associated with angiogenesis inhibitors (13).

VEGF belongs to a multigene family comprised of five members that bind to and selectively activate several membrane-bound tyrosine kinase receptors (VEGFR-1, -2, and -3) and neuropilins (NRP-1 and -2). VEGF (also known as VEGF-A) induces a robust proliferative response in endothelial cells by its interaction with VEGFR-2. The axis VEGF-A/VEGFR-2 is believed to be a key pathway in the angiogenic process and is the focus of current VEGF-based therapies. Moreover, other family members, such as placenta growth factor (PlGF), VEGF-B, and the receptor VEGFR-1, were initially met with less enthusiasm as potential targets, because their roles in angiogenesis were unclear; however, this view has been challenged by the development of compounds targeting these molecules. VEGFR-1 and NRP-1, as well as their specific ligands VEGF-B and PlGF, have a prominent role in angiogenesis and are promising therapeutic agents (14–20). Monoclonal antibodies directed against VEGFR-1, NRP-1, or PlGF have also shown potential as antitumor agents (18, 19, 21). In light of these studies and based on our own previous work (22), we reasoned that small molecules targeting the VEGF-receptor

---

pathways would potentially inhibit angiogenesis and therefore, would be prospects for future clinical studies.

Results and Discussion

Rational Small-Molecule Design. We have previously used a subtractive phage display-library screening on activated human endothelial cells to isolate a VEGFR-1– and NRP-1–binding peptide, CPQPRPLC (22). Here, to gain insight into the specific ligand-receptor requirements, we generated a panel of mutant peptide-targeted phage clones for use in structure-functional analysis. Unexpectedly, this in tandem strategy of peptide alanine scan followed by binding assays, clearly failed to identify the individual residues in CPQPRPLC that interact with the VEGF receptors (Fig. S1), indicating that methods other than site-directed mutagenesis would be required to unveil the fine molecular basis of this receptor-ligand interaction. We assigned NMR to the minimal structural requirements for binding of CPQPRPLC to VEGFR-1 and NRP-1 to the Arg_5-Pro_6-Leu_7 motif (RPL) embedded within the full peptide (23). Because it is established that peptide sequences selected by phage display often bind to biologically active sites in proteins (24, 25), we hypothesized that the tripeptide RPL, by its targeting of the extracellular domain of VEGFR-1 and NRP-1, might comprise the essential structure for a class of small tyrosin-kinase receptor inhibitors with antiangiogenic properties.

To develop our working hypothesis, we had to overcome a general limitation: peptides are generally considered notoriously labile and are too metabolically unstable to be suitable as drugs. We, therefore, synthesized peptidomimetic derivatives based on this VEGFR-1– and NRP-1–binding motif, and we tested their effects in angiogenesis assays in vivo.

In the design of peptidomimetics, we used a method known as amino acid retro-inversion (26, 27), which consists of the reversal of the peptide backbone stereochemistry (i.e., substitution of D-amino acids for the normal L-amino acids) in conjunction with chain reversal. The resulting product is a "retro-inverso" peptidomimetic in which the side-chain topology is similar to the parent peptide (Fig. 1A). Because mammalian enzymes do not effectively recognize D-amino acid residues and their bonds, these compounds are usually less susceptible to proteolytic degradation. In this context, the corresponding RPL peptide retro-inverted ᴅLeu-ᴅProᴅArg sequence, ᴅ(LPR) peptidomimetic, was chemically produced in solid phase as a lead compound for these studies. We subsequently incubated both RPL and ᴅ(LPR) with increasing concentrations of pancreatin, and the reaction products were analyzed by mass spectrometry. No degradation of ᴅ(LPR) was observed at the highest ratio of enzyme:peptide concentration (400 pg/nmol) tested (Fig. 1B); in contrast, RPL was markedly degraded by proteases, a conclusion based on the presence of a Pro-Leu fragment and the decreased intensity of the RPL peak relative to the peak of its degradation product (Fig. 1C). We used pancreatin, because it is a harsh proteolytic mixture of digestive enzymes including amylases, lipases, and proteases (such as trypsin and chymotrypsin); we reasoned that degradation resistance to this protease cocktail would strengthen our working hypothesis that a retro-inverted sequence results in a more stable drug. However, as a cautionary note, it has long been determined that D-amino acid oxidase is the only known mammalian enzyme that metabolize D-peptidomimetics in the kidney (28). Thus, despite the apparent resistance of ᴅ(LPR) to degradation by pancreatin, detailed pharmacological studies will be required to determine whether or not gastrointestinal administration in preclinical models or ultimately, oral use in patients will be a future possibility.

VEGF Receptor-Ligand Binding and Structure–Function Relationship. To empirically assess whether or not retro-inversion yielded an agent that binds to VEGFR-1 and NRP-1, we performed competition experiments. In this assay, VEGFR-binding phage particles (displaying CPQPRPLC) (22, 23) are incubated with receptors

(VEGFR-1 or NRP-1) in the presence of increasing concentrations of RPL or ᴅ(LPR). Binding of the RPL peptide or its corresponding peptidomimetic to the immobilized VEGFR-1 or NRP-1 receptor would compete with the binding of CPQPRPLC phage particles, the affinity of which for VEGFR-1 or NRP-1 has been calculated (22). The concentration of ligand required to inhibit 50% of CPQPRPLC phage binding to the receptor (IC_{50}) is then calculated for RPL and ᴅ(LPR). These competition experiments showed that both RPL and ᴅ(LPR) inhibited, in a concentration-dependent manner, the binding of CPQPRPLC phage to either receptor (Fig. 2A), whereas a control peptide used at the highest concentration (100 μM) had no effect on phage binding to VEGFR-1 or NRP-1. The IC_{50} of RPL for VEGFR-1 (30 nM) and NRP-1 (4 pM) and the IC_{50} of ᴅ(LPR) for VEGFR-1 (2 pM) and NRP-1 (2 pM) were reproducibly determined.

To understand the ligand attributes of our lead compound, we analyzed the direct interaction of ᴅ(LPR) to VEGF receptors by NMR, the methodology originally used to yield the functional RPL motif (23). Here, the VEGF receptor-ligand behavior of ᴅ(LPR) was first analyzed by total correlation spectroscopy (TOCSY) and nuclear overhauser enhancement apectroscopy (NOESY), and all resonances were unambiguously assigned (Table S1). To further identify the specific residues in ᴅ(LPR) that participate in the binding to either VEGFR-1 or NRP-1, we next incubated ᴅ(LPR) with each individual receptor and measured changes in NMR parameters (chemical shift changes) of ᴅ(LPR) resonances in the fast-exchange regime (the resonances in the peptidic NMR spectrum reflect the average parameters between the free and bound states). We observed that both VEGFR induced side-chain chemical shift changes in the resonances of the peptidomimetic (Fig. 2B and Table S2). Leu and Arg side-chain hydrogen atoms showed chemical shift changes in several residues (Table S2) as a consequence of incubation with VEGFR-1 and
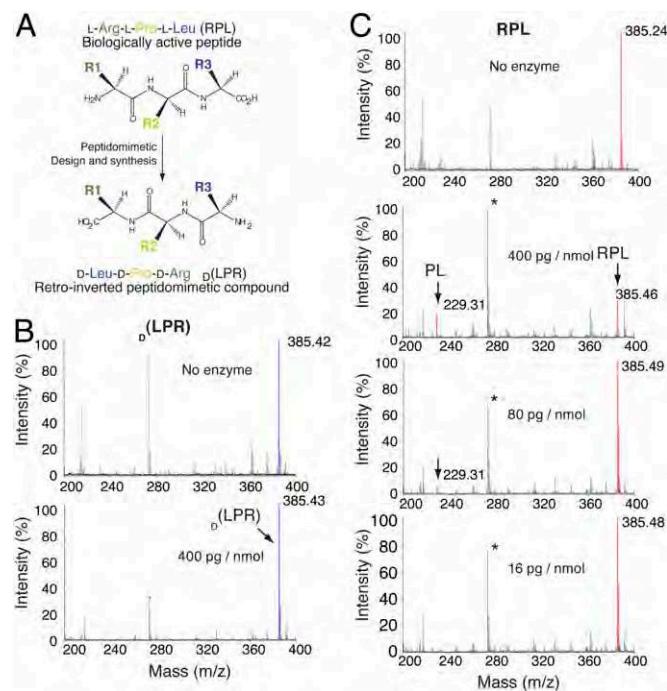
Fig. 1. Drug design and protease degradation-resistance assay. (A) Schematic representation of Arg-Pro-Leu (RPL) retro-inversion. (B and C) Mass-spectroscopy analysis (MALDI-TOF) of RPL and ᴅ(LPR) pre- and postincubation with pancreatic enzymes. Peaks corresponding to the intact peptides and enzymatic degradation products are color-coded [red, RPL; blue, ᴅ(LPR)] and indicated by arrows.
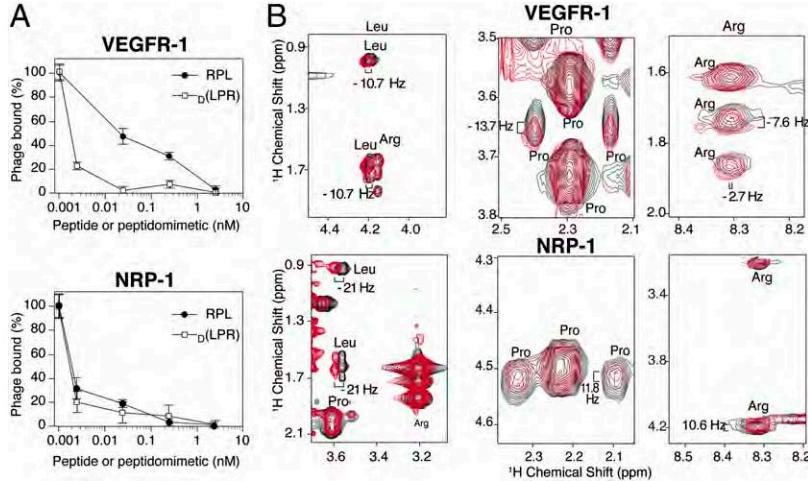
Fig. 2. The tripeptide RPL and the drug $_D$(LPR) target VEGFR-1 and NRP-1. (*A*) In a phage-competition assay, increasing concentrations of RPL (black circles) or $_D$(LPR) (open squares) inhibit binding of CPQPRPLC-displaying phage to the immobilized receptors VEGFR-1 or NRP-1. (*B*) Chemical-shift changes induced on the $_D$(LPR) peptidomimetic resonances by its binding to VEGFR-1 or NRP-1 at 25 °C are shown. 2D TOCSY spectra of $_D$(LPR) alone (black color) or in the presence of the individual receptors (red color) are shown. Different regions of the spectra of $_D$(LPR) are shown to indicate chemical-shift changes in individual $_D$(LPR) residues.

NPR-1, a result consistent with our studies of the RPL motif (23). In previous work, we showed that the interaction of RPL with NRP-1, but not with VEGFR-1, requires the key participation of the Pro residue, likely representing an evolutionary gain-of-function mutation (23). However, differently from the RPL motif, the Pro residue in $_D$(LPR) also participates in the binding to VEGFR-1, which is shown by the chemical shift change of 13.7 Hz in the Pro.HD1 hydrogen. These results confirm that $_D$(LPR) binds to VEGFR-1 and NRP-1, but, in contrast to RPL, binding of $_D$(LPR) to each individual receptor involves all amino acids in the peptidomimetic, including the Pro residue; these data could explain the similar affinity of $_D$(LPR) observed for either VEGF receptor. Consistently, a differential binding pattern to each VEGF receptor was not observed with the $_D$(LPR) peptidomimetic, which has a similar IC_{50} for either VEGFR-1 or NRP-1.

Together, we show (*i*) that the mimic $_D$(LPR) is less prone to proteolytic degradation, (*ii*) that both RPL and $_D$(LPR) have a higher affinity for NRP-1, but not for VEGFR-1, than that observed with the original CPQPRPLC (22), and (*iii*) that retro-inversion confers superior ligand binding to VEGF receptors; indeed, the IC_{50} of CPQPRPLC for NRP-1 (~50–100 nM) (22) is ~1.2×10^4-fold to 5×10^4-fold higher than the IC_{50} of RPL and $_D$(LPR) for either receptor, a functional result suggesting that they are much stronger VEGF receptor ligands than the CPQPRPLC peptide.

Effect of $_D$(LPR) in Angiogenesis Models. Given its favorable affinity to the targeted VEGF receptors and resistance to degradation in vitro, we concluded that $_D$(LPR) is a suitable candidate for evaluation of angiogenesis assays in vivo. To that end, we used three standard animal models to evaluate the effect of $_D$(LPR) on angiogenesis (29–33).

First, immunocompetent mice were implanted s.c. with $VEGF_{165}$-containing Matrigel, with or without either $_D$(LPR) or a negative control. On postimplantation day 7, neovascularization within each Matrigel plug was determined by quantification of hemoglobin. Matrigel plugs that had been preloaded with $_D$(LPR) and $VEGF_{165}$ exhibited significantly fewer new vessels than Matrigel plugs that contained only $VEGF_{165}$; no detectable effect on blood-vessel formation was observed in the $VEGF_{165}$-containing Matrigel plugs plus a control peptidomimetic (Fig. 3*A*). These results indicate that $_D$(LPR) markedly inhibits VEGF-induced neovascularization in vivo.

Second, we checked for a systemic therapeutic effect in a SCID-mouse model of human angiogenesis in which human endothelial cells, cultured in vivo within polymer implants, grow to form a microvasculature; then, this merges with the host (i.e., mouse) capillaries. These functional neovessels are lined with human

endothelial cells that express angiogenesis markers and serve to transport the murine circulation. To evaluate the effect of $_D$(LPR), we maintained mice implanted with scaffolds containing human endothelial cells for 9 days. The endothelial cells formed non-functional tubular structures containing empty lumens that slowly matured to fully functional human blood vessels containing the murine hematic cells and other elements (30). The implanted mice received i.p. administration of $_D$(LPR) or control peptidomimetic daily on days 12–21. The scaffolds were removed, and the number of functional blood vessels was determined by Factor VIII immunostaining. On day 21, all animals developed functional blood vessels with the expected cell density; these chimeric blood vessels also expressed the angiogenesis marker von Willebrand factor (vWF), a receptor for Factor VIII (Fig. 3*B*). We observed a significant reduction (37%; *t* test with *P* < 0.001) in the number of blood vessels in mice treated with $_D$(LPR) [32.1 ± 3.8 per field in
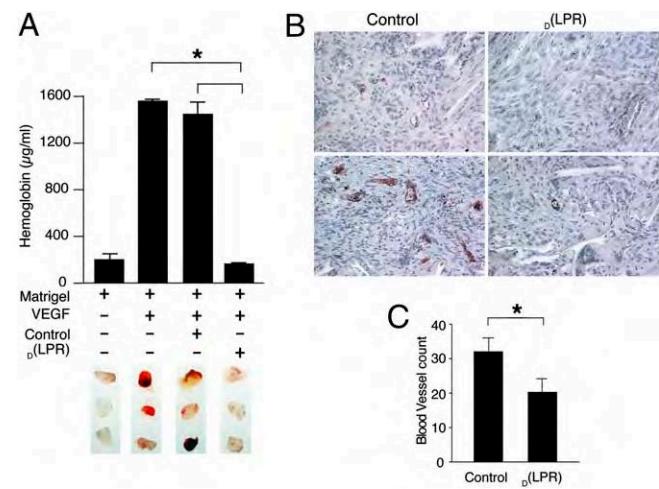
Fig. 3. Inhibition of neovascularization in vivo by $_D$(LPR) treatment. (*A*) Representative pictures of Matrigel plugs containing 500 μg/mL of $_D$(LPR) or control after 7 days of implantation (*Lower*). Matrigel plugs were excised, and angiogenesis was quantified by measurement of the hemoglobin content within the matrix. The bar graph shows representative mice from the experiment (*Upper*). (*B*) Immunostaining with anti-human factor VIII antibody of scaffolds containing human microvascular endothelial cells (HDMEC) implanted into SCID mice that received 25 mg/kg i.p. daily of $_D$(LPR) or control. (*C*) Number of human factor VIII-positive blood vessels at 200× magnification. *Student's *t* test (*P* < 0.01).

the control peptidomimetic versus 20.3 ± 3.8 in the presence of D(LPR)]. A typical experiment is shown in Fig. 3.

Third, we used a mouse model of retinopathy of prematurity (3, 31–34). This model is generally accepted not only as representative of the human condition, but also as the best available model of selected features of other retinal angiogenic diseases, including diabetic retinopathy and perhaps, age-related macular degeneration (3, 33). In brief, mice are exposed to 75% O$_2$ from postnatal day 7 (P7) to P12, a treatment that inhibits the formation of retinal blood vessels that are otherwise a normal developmental process at this age. The pups are returned to room air (21% O$_2$). Their retinas are analyzed at P17–P21, by which time pathological neovascularization has supervened on the inner retinal surface (31–33). On their return to room air (P12), neonatal mice received either D(LPR) or control daily for 7 days; another negative control cohort of mice received daily i.p. administration of the vehicle alone (PBS). At the end of treatment (P19), the eyes were examined histologically, and neovascularization was quantified by counting the number of blood vessels and endothelial cells protruding from the retina; a typical experiment is shown (Fig. 4A). Overall, a significant reduction in angiogenesis (59%; t test with $P < 0.05$) was observed in D(LPR)-treated mice ($n = 7$; median = 12; range = 0–5) relative to mice treated with the negative control ($n = 8$; median 26; range = 12–37) or the vehicle ($n = 8$; median = 29; range = 15–43) alone (Fig. 4B). We also used this model to compare the efficacy of D(LPR) with bevacizumab, a human monoclonal antibody. One should note that although bevacizumab has been reported to have lower (or no) reactivity with murine than with human VEGF (35), it has still shown antiangiogenic activity in ocular and tumor mouse models (36) and has later served as a positive control for new compounds in other experimental settings (37). At P12, mice received D(LPR) daily (20 mg/kg per day) or bevacizumab every other day (1 mg/kg per dose). A significant reduction in angiogenesis (72%; t test with $P < 0.05$) was observed in either D(LPR)-treated mice ($n = 9$; median = 11; range = 4–26) or bevacizumab-treated mice ($n = 9$; median = 12; range = 1–35) relative to mice that received vehicle alone ($n = 4$; median = 42; range = 19–70); this direct comparison in a side-by-side experiment (Fig. 4C) does suggest that the magnitude of the systemic therapeutic effect of D(LPR) may be in a similar range of other FDA-approved agents in clinical use.

The recent report, in which the inhibitor of *Src* and *Yes* administered topically blocked pathological angiogenesis in animal models (37), stimulated our reasoning to determine if D(LPR) would also penetrate the vitreous humor and be effective as an eye-drop (e.d.) formulation for topical application (Fig. 4D and Fig. S2 A–C). We first found that topically administered D(LPR) to the eyes of mice resulted in accumulation of the drug ($n = 6$; median = 1.25 μM; range = 0.65–2.14 μM) within the vitreous humor (Fig. S2 A–C). Then, to evaluate the effect of topically administered D(LPR), we put P7 mice in the oxygen chamber; at P12, animals were treated by laying D(LPR)-containing eye drops on the cornea (200 μg e.d. three times a day) for 7 days. At the end of treatment on P19, the eyes were examined histologically, and neovascularization was quantified. Again, despite minimal optimization in dose and schedule, a significant reduction in angiogenesis (53%; t test with $P < 0.05$) was observed in mice treated topically with D(LPR) ($n = 8$; median = 7; range = 2–23) relative to those receiving vehicle (Systane) alone ($n = 9$; median = 16; range = 5–43; Fig. 4D). These results suggest that D(LPR) may serve as a lead for the development of soluble and permeable small drug molecules that can be administered in eye drops. If so, an eye-drop drug formulation would represent a great improvement for many patients with blinding diseases who now must receive intraocular injections of agents with larger molecular weights, such as monoclonal antibodies.

Finally, in preliminary studies, we have begun to evaluate whether or not D(LPR) might also have effects against tumors. As an initial tumor model, we have used an isogenic mouse mammary
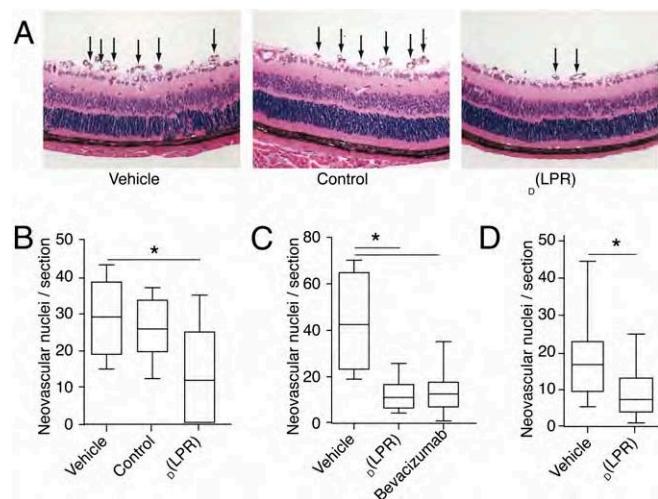
Fig. 4. Systemic and topical treatment with D(LPR) inhibits retinal angiogenesis in a mouse retinopathy-of-prematurity model. Retinal neovascularization was induced in C57BL/6 neonatal mice by exposure to 75% oxygen (P7–P12) followed by daily i.p. administration of D(LPR) or control (20 mg/kg per day). (A) H&E-stained retinal sections (P19) showed new blood vessels at the retinal inner surface (arrows) in control animals (*Top* and *Center*). There was marked reduction in D(LPR)-treated mice (*Bottom*). (B) P19 quantification of neovascular nuclei protruding into the vitreous space. Serial sections ($n > 5$) of eyes ($n > 10$) were quantified in each group. Treatment with D(LPR) yielded a significant reduction in nuclei relative to vehicle or control. (C) In systemic (i.p.) delivery, the magnitude of D(LPR)-induced neovascularization inhibition was similar to that observed by treatment with the monoclonal antibody bevacizumab. (D) Topical delivery of D(LPR) in an eye-drop formulation (200 μg e.d. three times daily) also induced a significant reduction in abnormal retinal angiogenesis. Shown is mean ± SEM in each experiment.*, Student's t test ($P < 0.05$).

cancer that is both aggressive and angiogenic (38). We first confirmed the ability of the CPQPRPLC-targeted phage to home i.v. to established tumors in vivo relative to the control nontargeted phage (Fig. 5A). We next treated tumor-bearing mice with vehicle, control peptide, or D(LPR). A significant reduction in tumor volume was detected in the cohort receiving D(LPR) relative to the control cohorts (Fig. 5B). Future preclinical studies will determine the full translational value of this drug prototype against cancer in addition to angiogenic retinopathies.

Mechanism of Action of D(LPR) on VEGF Receptor Pathways. The well-established involvement of VEGFR-1 and NRP-1 in endothelial cell proliferation (15, 39) has prompted us to evaluate the effects of the drug D(LPR) in this setting. However, several experimental challenges are evident in this complex ligand-receptor systematic analysis. VEGF$_{165}$ also binds to VEGFR-2, an interaction that mediates most of its mitogenic activity; moreover, PlGF, a VEGFR-1– and NRP-1–specific ligand that induces VEGFR-1 phosphorylation, extracellular signal-regulated kinase (ERK) activation, and endothelial cell proliferation (15, 39), is another potential target in angiogenesis (13, 21). Thus, given the similar receptor-binding profiles of the RPL motif and PlGF (22, 23), we chose to first study the effect of D(LPR) on this specific pathway. Human umbilical-vein endothelial cells (HUVEC) incubated with PlGF in the presence of increasing concentrations of D(LPR) showed a marked growth reduction relative to HUVEC cultured with no D(LPR); a control peptide had no detectable effect (Fig. 6A). Consistently, in contrast to its effect on PlGF-induced HUVEC proliferation, D(LPR) had no detectable effect on cell growth in the presence of VEGF$_{165}$, which acts mainly through the VEGFR-2 pathway.

To evaluate whether or not the inhibition of endothelial cell proliferation by D(LPR) is mediated through its binding to VEGFR-
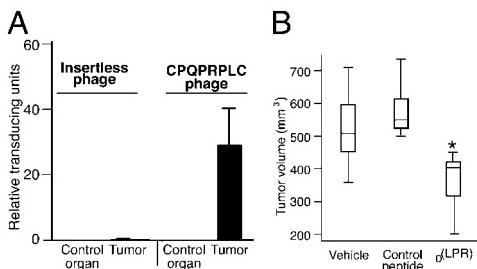
Fig. 5. CPQPRPLC-targeted phage homed to tumors, and D(LPR) treat-ment reduced tumor growth. (*A*) Tumor-bearing mice received 10^{10} transducing units (TU) of either CPQPRPLC phage or insertless phage (negative control). After 24 h of circulation, phage homing to tissues was evaluated by counting the relative TU. The brain served as a negative control organ. (*B*) Tumor-bearing mice ($n = 7$ per cohort) were treated with vehicle alone, D(LPR), or control peptide (50 mg/kg/day for 5 days). Two independent experiments were performed with similar results. A representative experiment is shown. *Student's *t* test ($P = 0.02$).

1 and/or NRP-1, we analyzed the phosphorylation of VEGFR-1 and ERK. D(LPR) inhibited VEGFR-1 Tyr1213 phosphorylation and reduced ERK phosphorylation by ~55% (Fig. 6*B*; Fig. S3 *A* and *B*), again in agreement with the previous results (Fig. 6*A*). VEGFR-1 and ERK phosphorylation were not altered by a negative control peptide. Similarly, we observed that the treatment of cells with D(LPR) affected neither PlGF-induced Akt phosphorylation nor ERK phosphorylation if endothelial cells were stimulated by VEGF instead of PlGF; both of these results suggest specificity. D(LPR) behaved similarly to other NRP-1–binding peptides with anti-angiogenic activity. Tuftsin (TKPPR) (40) and hexapeptide A7R (ATWLPPR) (41) both bind to NRP-1 and inhibit neo-vascularization in different assays; another peptide derived from the VEGF-E C-terminal domain (RPPR) also binds to NRP-1 and inhibits angiogenesis (42). Of note, all of these NRP-1–binding sequences share the terminal dipeptide motif Pro–Arg. Recently, a consensus sequence $^R/_KXX^R/_K$ has also been reported as an NRP-1–binding motif (43). Because D(LPR) does not functionally affect VEGF$_{165}$-stimulated endothelial cell activation and proliferation through VEGFR-2, we conclude that D(LPR) likely exerts its anti-angiogenic effects through VEGFR-1– and/or NRP-1–specific pathways; to our knowledge, RPL and D(LPR) are the only known reagents that target both NRP-1 and VEGFR-1. In summary, these signal-transduction data establish the basis of a specific mechanism of action for the RPL motif and its degradation-resistant counter-part, the peptidomimetic D(LPR).
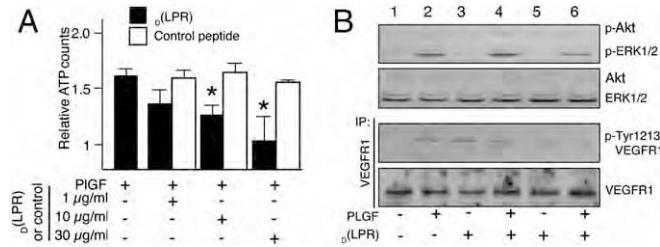
Fig. 6. Effect of D(LPR) on VEGFR-1–mediated endothelial cell pro-liferation and signaling. (*A*) Dose-dependent effect of D(LPR) on PlGF-induced HUVEC proliferation. Results are presented as values relative to BSA. (*B*) Immunoblot analysis of phosphorylated and total forms of VEGFR-1, ERK, and Akt. P1GF was used at 100 ng/mL, and D(LPR) or control was used at 10 µg/mL.

Conclusion

Based on a subtractive combinatorial screening on human angio-genic endothelial cells and structure-function studies (22, 23), here we rationally designed, synthesized, and functionally validated small-molecule targeting VEGF receptors, namely VEGFR-1– and NRP-1–specific pathways. This antiangiogenic drug, D(LPR), showed promising attributes in vitro, in cellulo, and in vivo in several mouse models. The data presented in this study support the working hypothesis that drugs against members of the VEGF family that do not signal through the VEGFR-2 axis may also have important biomedical applications. This fact reinforces recent reports of the roles of VEGFR-1 and NRP-1 in abnormal angio-genesis (18, 19, 44).

Currently available inhibitors of receptor tyrosine-kinase recep-tors belong to one of two classes of drugs: small molecules affecting the intracellular kinase domain or large antibodies binding to the extracellular domain. Our results show that small molecules tar-geting an extracellular receptor ligand-binding domain comprise an unrecognized candidate class of receptor tyrosine-kinase inhibitors.

Even without much optimization in dose and schedule, we show that D(LPR) is water-soluble and degradation-resistant, which are favorable biochemical features for possible translation into an orally or topically administered drug. Of course, although small molecules might have a few inherent theoretical advantages with respect to tissue permeability, biodistribution, and cost effective-ness, it remains to be determined whether or not the prototype introduced here will prove at least as effective as currently avail-able drugs (7–10) or newly engineered antibodies (45). If so, development of D(LPR) or its derivatives as antiangiogenic drugs in an eye-drop formulation would represent an improvement in quality of life for patients with retinopathies. Unequivocally, using eye drops to avoid distinctly unpleasant, risky, and expensive repetitive injections into the eye would have universal appeal.

Material and Methods

Reagents and Cell Culture. All peptides and peptidomimetics were synthesized and purified by HPLC to a purity greater than 95% by Polypeptide Labo-ratories. The parental peptide sequence CPQPRPLC, along with its motifs [i.e., linear or cyclic RPL peptide, linear or cyclic negative control Ala-Pro-Ala (APA) peptide, linear or cyclic D(LPR) peptidomimetic, and linear or cyclic negative control peptidomimetic D(APA)], were used correspondingly as appropriate, unless otherwise specified.

VEGF receptors and ligands were purchased from R&D Systems or Leinco Technologies. Heparin, Drabkin's reagent, hemoglobin, and Brij-35 were obtained from Sigma-Aldrich. Phospho-tyrosine monoclonal antibody (sc-7020) was purchased from Santa Cruz Biotechnology, and VEGFR-1 poly-clonal sera were purchased from R&D Systems (AF321). Anti-pERK1/2, pAkt (Ser473), anti-Akt, and anti-p44/42 MAPK (ERK1/2) were obtained from Cell Signaling Technology. HUVEC (Lonza) was grown on gelatin-coated dishes in endothelial-cell basal medium (EBM) containing endothelial growth factors (Lonza), 2% heat-inactivated FBS, 50 U/mL penicillin-streptomycin, and 2 mM L-glutamine. Cells were maintained at 37 °C and 5% CO_2.

Animals. The Institutional Animal Care and Use Committees at the University of Texas M.D. Anderson Cancer Center and the University of Michigan approved all animal experimentation. This study adhered to the Association for Research in Vision and Ophthalmology (ARVO) statement for the use of animals in ophthalmic and vision research. Mice were bought from Harlan.

Protease Degradation-Resistance Assay. D(LPR) or RPL was diluted to 500 µg/mL in PBS and incubated with serially increasing concentrations of pancreatin (Sigma-Aldrich) for 2 h at 37 °C. Samples were analyzed by MALDI-TOF.

Peptide-Targeted Phage Binding and Ligand-Competition Assay. A phage-binding assay on purified proteins and ligand-competition assays were car-ried out as described (22) in *SI Materials and Methods*.

NMR Spectroscopy. NMR experiments were performed as described (23, 46) in *SI Materials and Methods*.

Cell Proliferation, Signal Transduction, and Phosphorylation Analysis. We analyzed HUVEC proliferation, signal-transduction pathways, and VEGF receptor phosphorylation; for details, see *SI Materials and Methods*.

Mass Spectrometry Analysis of Drug Penetration. $_D$(LPR) was detected and quantified in the vitreous humor after topical e.d. application. $_D$(LPR) (2 μL of a 100 mg/mL solution in Systane) was dropped topically on the external corneal surface of C57BL/6 mice (e.d. every 2 h; $n = 3$ doses). Two hours after the third and final dose, mice were killed, eyes were enucleated, and vitreous humor was gently released and collected by centrifugation (10,000 × g for 1 minute). Vitreous humor samples or $_D$(LPR) standards were diluted 1:10 in α-cyano-4-hydroxyl cinnamic acid (10 mg/mL in 50:50 acetonitrile:water; 0.1% trifluoro-acetic acid final concentration), spotted on the target, and analyzed in a MALDI-TOF/TOF System [4700 Proteomics Analyzer; Applied Biosystems (ABI)].

Angiogenesis Assays. We used an in vivo Matrigel angiogenesis assay as described (29). For the model of human angiogenesis, SCID mice were implanted with 10^6 human dermal microvascular endothelial cells (HDMEC) highly porous poly-L (lactic) acid scaffolds. Two scaffolds were used per mouse, one on each flank (30). For the retinal neovascularization angiogenesis assay, we used C57BL/6 mouse pups with their nursing mothers (33). For details, see *SI Materials and Methods*.

Tumor Targeting. Selective phage homing to tumors and peptide treatment were performed as described (38). See *SI Materials and Methods* for details.

Statistics. The appropriate statistical test was used for the analysis of assays as indicated. Statistical significance was set at $P < 0.05$. For details, see *SI Materials and Methods*.

1. Folkman J (1971) Tumor angiogenesis: Therapeutic implications. *N Engl J Med* 285:1182–1186.
2. Folkman J (2007) Angiogenesis: An organizing principle for drug discovery? *Nat Rev Drug Discov* 6:273–286.
3. Arden GB, Sidman RL, Arap W, Schlingemann RO (2005) Spare the rod and spoil the eye. *Br J Ophthalmol* 89:764–769.
4. Carmeliet P (2005) Angiogenesis in life, disease and medicine. *Nature* 438:932–936.
5. Loges S, Roncal C, Carmeliet P (2009) Development of targeted angiogenic medicine. *J Thromb Haemost* 7:21–33.
6. Adams RH, Alitalo K (2007) Molecular regulation of angiogenesis and lymphangio genesis. *Nat Rev Mol Cell Biol* 8:464–478.
7. Hurwitz H, et al. (2004) Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med* 350:2335–2342.
8. Crawford Y, Ferrara N (2009) VEGF inhibition: Insights from preclinical and clinical studies. *Cell Tissue Res* 335:261–269.
9. Gragoudas ES, et al. (2004) Pegaptanib for neovascular age-related macular degeneration. *N Engl J Med* 351:2805–2816.
10. Rosenfeld PJ, et al. (2006) Ranibizumab for neovascular age-related macular degeneration. *N Engl J Med* 355:1419–1431.
11. Ellis LM, Hicklin DJ (2008) VEGF-targeted therapy: Mechanisms of anti-tumour activity. *Nat Rev Cancer* 8:579–591.
12. Ellis LM, Hicklin DJ (2008) Pathways mediating resistance to vascular endothelial growth factor-targeted therapy. *Clin Cancer Res* 14:6371–6375.
13. Fischer C, Mazzone M, Jonckx B, Carmeliet P (2008) FLT1 and its ligands VEGFB and PlGF: Drug targets for anti-angiogenic therapy? *Nat Rev Cancer* 8:942–956.
14. Carmeliet P, et al. (2001) Synergism between vascular endothelial growth factor and placental growth factor contributes to angiogenesis and plasma extravasation in pathological conditions. *Nat Med* 7:575–583.
15. Autiero M, et al. (2003) Role of PlGF in the intra- and intermolecular cross talk between the VEGF receptors Flt1 and Flk1. *Nat Med* 9:936–943.
16. Luttun A, et al. (2002) Revascularization of ischemic tissues by PlGF treatment, and inhibition of tumor angiogenesis, arthritis and atherosclerosis by anti-Flt1. *Nat Med* 8:831–840.
17. Kaplan RN, et al. (2005) VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. *Nature* 438:820–827.
18. Wu Y, et al. (2006) Anti-vascular endothelial growth factor receptor-1 antagonist antibody as a therapeutic agent for cancer. *Clin Cancer Res* 12:6573–6584.
19. Pan Q, et al. (2007) Blocking neuropilin-1 function has an additive effect with anti-VEGF to inhibit tumor growth. *Cancer Cell* 11:53–67.
20. Zhang F, et al. (2009) VEGF-B is dispensable for blood vessel growth but critical for their survival, and VEGF-B targeting inhibits pathological angiogenesis. *Proc Natl Acad Sci USA* 106:6152–6157.
21. Fischer C, et al. (2007) Anti-PlGF inhibits growth of VEGF(R)-inhibitor-resistant tumors without affecting healthy vessels. *Cell* 131:463–475.
22. Giordano RJ, Cardó-Vila M, Lahdenranta J, Pasqualini R, Arap W (2001) Biopanning and rapid analysis of selective interactive ligands. *Nat Med* 7:1249–1253.
23. Giordano RJ, et al. (2005) Structural basis for the interaction of a vascular endothelial growth factor mimic peptide motif and its corresponding receptors. *Chem Biol* 12:1075–1083.
24. Falciani C, Lozzi L, Pini A, Bracci L (2005) Bioactive peptides from libraries. *Chem Biol* 12:417–426.
25. Sergeeva A, Kolonin MG, Molldrem JJ, Pasqualini R, Arap W (2006) Display technologies: Application for the discovery of drug and gene delivery agents. *Adv Drug Deliv Rev* 58:1622–1654.
26. Chorev M (2005) The partial retro-inverso modification: A road traveled together. *Biopolymers* 80:67–84.
27. Fischer PM (2003) The design, synthesis and application of stereochemical and directional peptide isomers: A critical review. *Curr Protein Pept Sci* 4:339–356.
28. Meister A (1965) *Biochemistry of the Amino Acids* (Academic, New York), 2nd Ed, Vol 1, pp 364.
29. Passaniti A, et al. (1992) A simple, quantitative method for assessing angiogenesis and antiangiogenic agents using reconstituted basement membrane, heparin, and fibroblast growth factor. *Lab Invest* 67:519–528.
30. Nör JE, et al. (2001) Engineering and characterization of functional human microvessels in immunodeficient mice. *Lab Invest* 81:453–463.
31. Smith LE, et al. (1994) Oxygen-induced retinopathy in the mouse. *Invest Ophthalmol Vis Sci* 35:101–111.
32. Pierce EA, Avery RL, Foley ED, Aiello LP, Smith LE (1995) Vascular endothelial growth factor/vascular permeability factor expression in a mouse model of retinal neovascularization. *Proc Natl Acad Sci USA* 92:905–909.
33. Lahdenranta J, et al. (2001) An anti-angiogenic state in mice and humans with retinal photoreceptor cell degeneration. *Proc Natl Acad Sci USA* 98:10368–10373.
34. Heidary G, Vanderveen D, Smith LE (2009) Retinopathy of prematurity: Current concepts in molecular pathogenesis. *Semin Ophthalmol* 24:77–81.
35. Gerber HP, et al. (2007) Mice expressing a humanized form of VEGF-A may provide insights into the safety and efficacy of anti-VEGF antibodies. *Proc Natl Acad Sci USA* 104:3478–3483.
36. Dorrell MI, Aguilar E, Scheppke L, Barnett FH, Friedlander M (2007) Combination angiostatic therapy completely inhibits ocular and tumor angiogenesis. *Proc Natl Acad Sci USA* 104:967–972.
37. Scheppke L, et al. (2008) Retinal vascular permeability suppression by topical application of a novel VEGFR2/Src kinase inhibitor in mice and rabbits. *J Clin Invest* 118:2337–2346.
38. Hajitou A, et al. (2006) A hybrid vector for ligand-directed tumor targeting and molecular imaging. *Cell* 125:385–398.
39. Landgren E, Schiller P, Cao Y, Claesson-Welsh L (1998) Placenta growth factor stimulates MAP kinase and mitogenicity but not phospholipase C-gamma and migration of endothelial cells expressing Flt 1. *Oncogene* 16:359–367.
40. von Wronski MA, et al. (2006) Tuftsin binds neuropilin-1 through a sequence similar to that encoded by exon 8 of vascular endothelial growth factor. *J Biol Chem* 281:5702–5710.
41. Starzec A, et al. (2006) Antiangiogenic and antitumor activities of peptide inhibiting the vascular endothelial growth factor binding to neuropilin-1. *Life Sci* 79:2370–2381.
42. Cébe-Suarez S, et al. (2008) Orf virus VEGF-E NZ2 promotes paracellular NRP-1/VEGFR-2 coreceptor assembly via the peptide RPPR. *FASEB J* 22:3078–3086.
43. Teesalu T, Sugahara KN, Kotamraju VR, Ruoslahti E (2009) C-end rule peptides mediate neuropilin-1-dependent cell, vascular, and tissue penetration. *Proc Natl Acad Sci USA* 106:16157–16162.
44. Luttun A, Autiero M, Tjwa M, Carmeliet P (2004) Genetic dissection of tumor angiogenesis: Are PlGF and VEGFR-1 novel anti-cancer targets? *Biochim Biophys Acta* 1654:79–94.
45. Dennis MS, et al. (2007) Imaging tumors with an albumin-binding Fab, a novel tumor-targeting agent. *Cancer Res* 67:254–261.
46. Sklenar V, Piotto M, Leppik R, Saudek V (1993) Gradient-tailored water suppression for ^{1}H-15N HSQC experiments optimized to retain full sensitivity. *J Magn Reson Series A* 102:241–245.

Applying E_{max} model and bivariate thin plate splines to assess drug interactions

Maiying Kong[1] & J. Jack Lee[2]

[1]Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville, Louisville, Kentucky 40292, U.S.A.,[2]Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, Unit 1411, 1515 Holcombe Boulevard, Houston, Texas 77030, U.S.A.

TABLE OF CONTENTS

1. ABSTRACT

We review the semiparametric approach previously proposed by Kong and Lee and extend it to a case in which the dose-effect curves follow the *Emax* model instead of the median effect equation. When the maximum effects for the investigated drugs are different, we provide a procedure to obtain the additive effect based on the Loewe additivity model. Then, we apply a bivariate thin plate spline approach to estimate the effect beyond additivity along with its 95% point-wise confidence interval as well as its 95% simultaneous confidence interval for any combination dose. Thus, synergy, additivity, and antagonism can be identified. The advantages of the method are that it provides an overall assessment of the combination effect on the entire two-dimensional dose space spanned by the experimental doses, and it enables us to identify complex patterns of drug interaction in combination studies. In addition, this approach is robust to outliers. To illustrate this procedure, we analyzed data from two case studies.

2. INTRODUCTION

Studies of interactions among biologically active agents, such as drugs, carcinogens, or environmental pollutants have become increasingly important in many branches of biomedical research. For example, in cancer chemotherapy, the therapeutic effect of many anticancer drugs is limited when they are used as single drugs. Finding combination therapies with increased treatment effect and decreased toxicity is an active and promising research area (1). An effective and accurate evaluation of drug interaction for *in vitro* and/or *in vivo* studies can help to determine whether a combination therapy should be further investigated.

The literature supports the Loewe additivity model as the gold standard for defining drug interactions (2-5). The Loewe additivity model defines an additive effect based on the following equation

$$\frac{d_1}{D_{y,1}} + \frac{d_2}{D_{y,2}} = 1 \ . \qquad (E1)$$
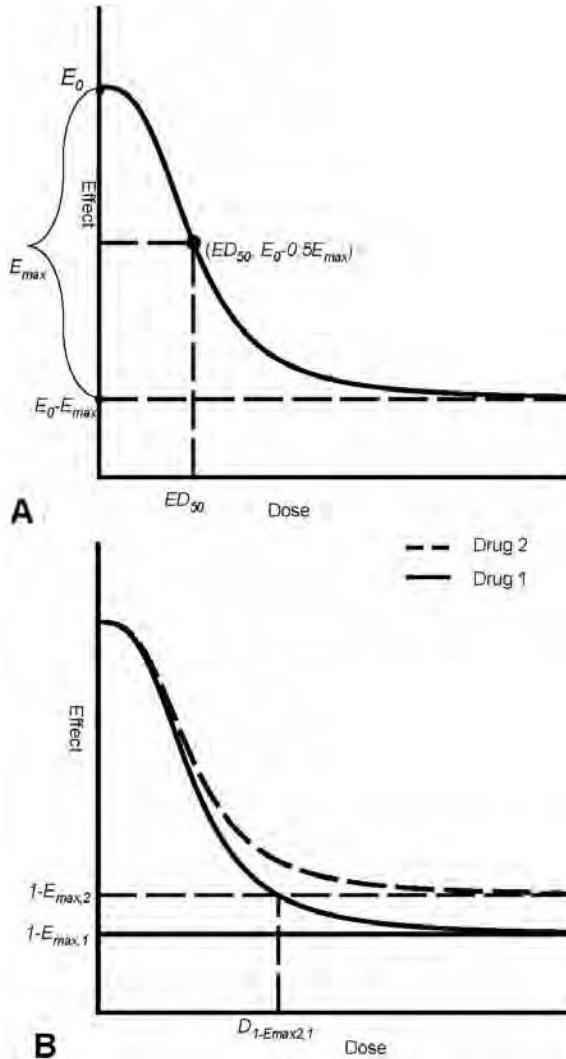
Figure 1. Dose-effect curves. Panel A shows a typical curve with the maximum effect, E_{max}, less than 1. ED_{50} is the dose required to produce half of the maximum effect, E_0-$0.5E_{max}$. Panel B shows two dose-effect curves with different maximum effects, say, $E_{max,1}$>$E_{max,2}$. In Panel B, drug 1 at dose level $D_{1\text{-}Emax2,1}$ produces the maximum effect produced by drug 2 alone.

Here y is the predicted additive effect, which is produced by the combination dose (d_1, d_2) when the two drugs do not interact; and $D_{y,1}$ and $D_{y,2}$ are the respective doses of drug 1 and drug 2 required to produce the same effect y when applied alone. If we know the dose-effect relationship for each single agent, say $E(d)=f_i(d)$ for agent i (i=1,2), we are able to obtain the dose $D_{y,i}$ by using the inverse function of f_i, denoted as $f_i^{-1}(y)$. By replacing $D_{y,1}$ and $D_{y,2}$ in equation ($E1$) with $f_1^{-1}(y)$ and $f_2^{-1}(y)$, respectively, we can obtain an equation that includes the single variable y, i.e.,

$$\frac{d_1}{f_1^{-1}(y)} + \frac{d_2}{f_2^{-1}(y)} = 1. \qquad (E2)$$

By solving equation ($E2$), we can obtain the predicted additive effect y. If the observed effect at (d_1,d_2) is more than (equal to, or less than) the predicted effect, we say that the combination dose (d_1, d_2) is synergistic (additive, or antagonistic).

In our previous studies (6-8), we found that Chou and Talalay's (9) median effect equation was appropriate to describe the dose-effect relationships. Chou and Talalay's median effect equation, in its nonlinear form, can be written as follows:

$$E = \frac{(d / ED_{50})^m}{1 + (d / ED_{50})^m} \qquad (E3)$$

where ED_{50} is the dose required to produce 50% of the maximum effect, and m is the slope factor (Hill coefficient), measuring the sensitivity of the effect to the dose range of the drug. For the data in the case studies (see Section 4 for details), we found that the median effect equation ($E3$) could not adequately describe the marginal dose-effect relationship because the plateau of the effect does not go to zero when a large dose level of a drug is applied. Instead, the following E_{max} model ($E4$) presented by Ting (10) describes the dose-effect relationship very well:

$$E = E_0 - \frac{(d / ED_{50})^m E_{max}}{1 + (d / ED_{50})^m}. \qquad (E4)$$

In the E_{max} model ($E4$), E_0 is the base effect, corresponding to the measurement of response when no drug is applied; E_{max} is the maximum effect attributable to the drug; ED_{50} is the dose level producing half of E_{max}, i.e., ED_{50} is the dose level required to produce the effect at a value of E_0-$0.5E_{max}$ (Figure 1.A); d is the dose level, which produces the effect E. Thus, E_0-E_{max} will be the asymptotic net effect when a large dose of the drug is applied. Different maximum effects for agents may reflect different mechanisms of action for the drugs (11). For *in vitro* studies, cell growth is commonly used as an endpoint to measure the effect of inhibitors. When no drugs (or, no inhibitors) are applied, the cell proliferation obtains its largest value. In this case, the dose-effect curve is similar to the one shown in Figure 1.A, where E_{max}>0. The effect range determined by the dose-effect curve lies between (E_0-E_{max}, E_0), and the asymptotic measurement for the maximum drug effect is E_0-E_{max}.

In the investigation of drug interactions, theoretically, we expect the measurements for the endpoints to be similar when no drug is applied. We use the measurements that are made without any drugs applied as controls. However, we realize that environmental factors other than the experimental conditions may lead to different measurements for the controls under different environments. Thus, we may need to standardize the observed effects by the mean of the control for each environmental condition (1, 6), and then take E_0=1. In this paper, we consider the following dose-effect curve for each drug:

$$E = 1 - \frac{(d / ED_{50})^m E_{max}}{1 + (d / ED_{50})^m}, \qquad (E5)$$

which assumes an effect at value 1 when no drug is applied. Once we obtain the dose-effect curve for each single drug, we can use the Loewe additivity model (*E1*) to obtain the additive effect for any combination dose, particularly, for the combination dose with observed effects. Thus, we may obtain the differences in observed effects and the predicted additive effect at each observed combination dose. We use the bivariate thin plate splines approach (12) to estimate the relationship between these differences and the combination doses. Consequently, we obtain a response surface of the differences over the combination doses, and can construct 95% confidence surfaces of the response surface. When the dose-response curves decrease with increasing dose, an observed effect that is smaller in magnitude than the prediction of Loewe additivity implies that the observed effect is stronger than the predicted effect, indicating that the combination dose is synergistic. Conversely, an observed effect that is larger in magnitude than the prediction of Loewe additivity implies that the observed effect is weaker than the predicted effect, indicating that the combination dose is antagonistic. However, these inferences should be made based on sound statistical considerations. Based on the fitted response surface and its upper and lower confidence surfaces, we can judge whether the difference is significantly less than zero, not different from zero, or greater than zero, and we can determine the patterns of drug interaction in terms of synergy, additivity, and antagonism. This paper is organized as follows. In Section 3.1, we describe the underlying stochastic assumption for the dose-effect curve and the procedure to estimate the parameters in each marginal dose-effect curve. In Section 3.2, we explain how we obtained the additive response surface based on the Loewe additivity model, in particular for studies in which the maximum effects of the drugs are different. In Section 3.3, we explain how we assessed the response surface beyond the additivity surface and how we constructed its 95% confidence surfaces. These procedures allow us to identify drug interactions in terms of synergy, additivity, or antagonism for all of the combination doses in the region containing the combination design points. In Section 4, we illustrate the procedure introduced in Section 3 by analyzing real data in two case studies. The last section is devoted to a short summary and perspective.

3. STATISTICAL METHODS

Assume that the observed data are (d_{1i}, d_{2i}, E_i) for $i=1, \ldots, n$. For each i, (d_{1i}, d_{2i}) is the observed combination dose and E_i is the corresponding observed effect. When only a single drug is applied (drug 1 or drug 2), we refer to the observations as marginal observations. That is, the marginal observations for drug 1 are the observations (d_{1i}, d_{2i}, E_i) with $d_{2i}=0$ ($i=1, \ldots, n$), and the marginal observations for drug 2 are the observations (d_{1i}, d_{2i}, E_i) with $d_{1i}=0$ ($i=1, \ldots, n$). The marginal dose-effect curves are estimated based on the marginal observations, which we present in Section 3.1. It is commonly accepted that the additive effect should be obtained based on the dose-effect relationships for each individual drug. In Section 3.2, we explain how we obtained the predicted

effect at combination dose (d_1, d_2) based on the Loewe additivity model (*E1*) and the marginal dose-effect curves (*E5*). We denote the predicted effect as $\hat{F}_p(d_1, d_2)$. By definition, there is no drug interaction when only a single drug is applied. Therefore, the term for drug interaction is meaningful only for the combination dose (d_1, d_2) with nonzero d_1 and d_2. In Section 3.3, we develop a procedure to estimate the effect beyond additivity for any combination dose (d_1, d_2) with nonzero d_1 and d_2, denoted by $\hat{f}(d_1, d_2)$.

3.1. Estimating dose-effect curves
Chou and Talalay (9), Chou (4), and Kong and Lee (6) estimated the parameters in the median effect equation (*E3*) by using the transformation $log\,E/(1-E) = m\,log(d/ED_{50}) = \alpha + m\,log(d)$ and applying the least squares method in the linear regression setting, where $\alpha = -m\,log(ED_{50})$. The case studies we evaluated (see Section 4) included many low doses, the effects of which are larger than 1 after setting the effect at the control level to be 1. Thus, a similar transformation for models (*E3*) and (*E5*) cannot be carried out. Because the measurements are continuous, we propose applying nonlinear least squares regression to estimate the parameters in models (*E3*) and (*E5*) with the assumption that a stochastic error with $N(0, \sigma^2)$ exists on the right-hand side of the two models. We note that estimating the dose-effect curve for drug i requires only the marginal observations for drug i with $i=1, 2$. We apply nonlinear least squares regression to estimate the parameters in the marginal dose-effect curves in the two case studies (shown in Section 4).

3.2. Predicting additive effects
We obtain the predicted effect based on the Loewe additivity model (*E 1*) when model (*E 5*) is applied as the marginal dose-effect curve for each drug. When model (*E 5*) is applied, the dose required to produce effect E is given by

$$d = ED_{50}\left(\frac{1-E}{E-(1-E_{max})}\right)^{1/m}.$$

However, the maximum effects for the two drugs may be different. Without a loss of generality, we assume that the maximum effect of drug 1 is larger than the maximum effect of drug 2, i.e., $E_{max,1} > E_{max,2}$. For this case, when the dose-effect curves are decreasing, neither drug applied alone can produce an effect in (0, $1 - E_{max,1}$) (Figure 1.B). Based on the Loewe additivity model (*E 1*), we can see that the predicted effect will be in the interval of ($1 - E_{max,1}$, 1) for any combination dose (d_1, d_2).

Recall that the Loewe additivity model (*E1*) can be rewritten as $d_1 + (D_{y,1}/D_{y,2})d_2 = D_{y,1}$, and the ratio $D_{y,1}/D_{y,2}$ (denoted as $\rho(y)$), is often called the relative potency of drug 2 versus drug 1 at effect level y, which means that the effect of 1 unit of drug 2 produces the same effect as $\rho(y)$ units of drug 1. Generally speaking, the relative potency $\rho(y)$ is dose-dependent (7). When there is no drug interaction, the effect of the combination dose (d_1, d_2) produces the same effect as drug 1 alone at dose level
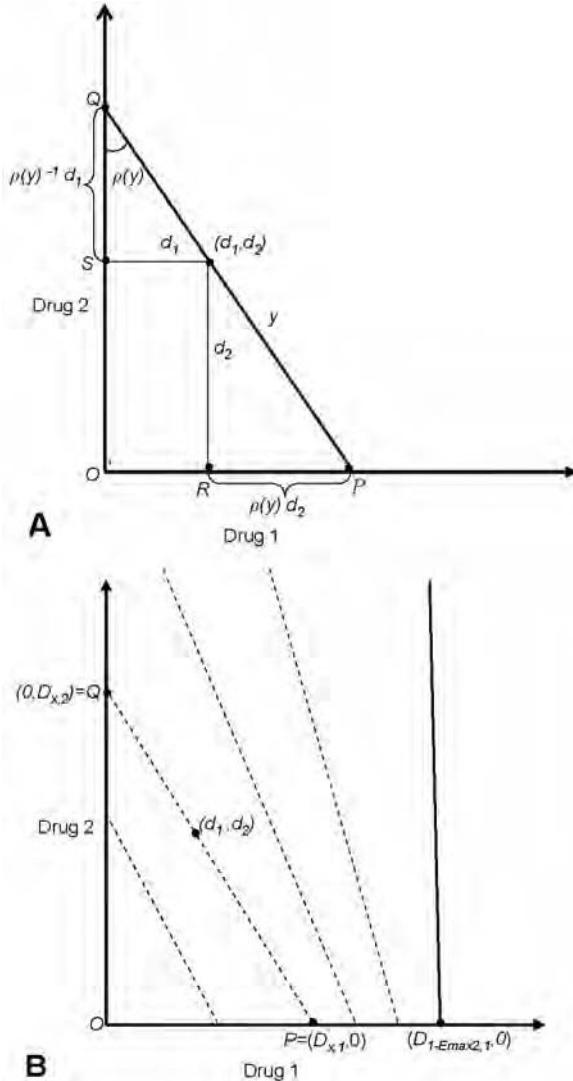
Figure 2. Additive isoboles. Panel A shows additive isobole under the Loewe additivity model. Any combination dose (d_1, d_2) on the line $\overline{PQ}$ produces the same effect as drug 1 alone at dose $D_{y,1}$ ($d_1 + \rho(y)d_2$), or drug 2 alone at dose $D_{y,2}$ ($\rho(y)^{-1}d_1 + d_2$), y is the predicted effect for any combination dose at the line $\overline{PQ}$, and $\rho(y)$ is the relative potency at the effect level y. Panel B shows that the additive isoboles associated with the effect level in $(1-E_{max2}, 1)$ cover the bound between the two solid vertical lines under the assumption $E_{max,1} > E_{max,2}$. Each dashed line corresponds to an isobole.

$D_{y,1}$, which equals $d_1 + \rho(y)d_2$, or drug 2 alone at dose $D_{y,2}$, which equals $\rho(y)^{-1}d_1 + d_2$ (Figure 2.A). All the combination doses (d_1, d_2) on the line $\overline{PQ}$ have the predicted effect y, where $\overline{PQ}$ is the line connecting the points P=($D_{y,1}$,0) and Q=(0,$D_{y,2}$) (Figure 2.A). This line $\overline{PQ}$ is often called an additive isobole (3, 4).

When $E_{max,1} > E_{max,2}$, as illustrated in Figure 1.B, we can calculate the dose of drug 1 required to produce the maximum effect of drug 2, i.e., $D_{1-E_{max,2},1} = ED_{50,1}\left(\dfrac{E_{max,2}}{E_{max,1}-E_{max,2}}\right)^{1/m_1}$. Note that the range of the effect for drug 2 is ($1 - E_{max,2}$, 1), which could be produced by drug 1 alone at a dose level between 0 and $D_{1-E_{max,2},1}$. Based on the Loewe additivity model, for any level of effect y in ($1-E_{max,2}$, 1), the associated additive isobole is the line connecting $(D_{y,1}, 0)$ and $(0, D_{y,2})$. When y varies from $1-E_{max,2}$ to 1, the dose of drug 1 required to produce effect y varies from $D_{1-E_{max,2},1}$ to 0, while the dose of drug 2 required to produce effect y varies from infinitely large to 0. In particular, when y is close to $1 - E_{max,2}$, the dose of drug 1 required to produce effect y is close to $D_{1-E_{max,2},1}$, and the dose of drug 2 required to produce effect y goes to infinity. Figure 2.B shows four typical additive isoboles (dashed lines), which connect equally effective doses of drug 1 and drug 2 at different effect levels. From left to right, the effect level decreases in magnitude. The additive isoboles may not be parallel because the relative potency may not be constant. When y varies in ($1-E_{max,2}$, 1), all the additive isoboles cover the region between the two solid vertical lines (Figure 2.B). Meanwhile, any combination dose (d_1, d_2) with $d_1 < D_{1-E_{max,2},1}$ must lie on one of these isoboles. Therefore, for any combination dose (d_1, d_2) with $d_1 < D_{1-E_{max,2},1}$, the predicted additive effect, say y, can be obtained by solving the following nonlinear equation for E:

$$\frac{d_1}{ED_{50,1}\left(\dfrac{1-E}{E-(1-E_{max,1})}\right)^{1/m_1}} + \frac{d_2}{ED_{50,2}\left(\dfrac{1-E}{E-(1-E_{max,2})}\right)^{1/m_2}} = 1.$$

Now we examine the predicted effect for the combination dose (d_1, d_2) with $d_1 \geq D_{1-E_{max,2},1}$. When $d_1 \geq D_{1-E_{max,2},1}$, drug 1 alone at dose d_1 produces an effect $E = 1 - \dfrac{(d_1/ED_{50,1})^{m_1} E_{max,1}}{1+(d_1/ED_{50,1})^{m_1}}$, an effect beyond $1 - E_{max,2}$, which cannot be produced by drug 2 alone at any dose level. In this case, if the effect of the combination dose is more than the effect produced by drug 1 alone, then drug 2 potentiates the effect of drug 1, and synergy occurs because the predicted additive effect is the effect produced by drug 1 alone at dose level d_1. Alternatively, because drug 2 alone cannot produce such an effect, we could consider $D_{y,2}$ to be infinitely large. Thus, the Loewe additivity model is reduced to $d_1/D_{y,1} = 1$, and the predicted additive effect is the effect produced by drug 1 alone at $D_{y,1}$. No matter which approach we take, the predicted effect y is the same. We can determine the predicted effect y from the following equation:

$$y = f_1(D_{y,1}) = f_1(d_1) = 1 - \frac{(d_1/ED_{50,1})^{m_1} E_{max,1}}{1+(d_1/ED_{50,1})^{m_1}}.$$

Thus, we can obtain the predicted effect for any combination dose (d_1,d_2). Using notation similar to that

from the previous study of Kong and Lee (6), we denote the predicted effect at the combination dose (d_1, d_2) as $\hat{F}_p(d_1, d_2)$. In the following subsection, we develop a procedure to estimate the effect beyond additivity, denoted by $\hat{f}(d_1, d_2)$ and to construct its 95% point-wise confidence interval and simultaneous confidence interval. We assess the drug-drug interaction based on the estimated $\hat{f}(d_1, d_2)$ and its confidence intervals.

3.3. Assessing drug interactions using bivariate thin plate splines

By definition, there is no drug interaction when a single drug is used alone. Therefore, we set the differences between the observed and predicted effects at zero for the marginal observations, that is, for the combination doses (d_1, d_2) with only one nonzero component. We apply a bivariate thin plate spline to estimate the differences as a function of the combination dose, say, $f(d_1, d_2)$. When the dose-effect curves are decreasing, $f(d_1, d_2) < 0$ indicates that the observed effect is more than the predicted effect at (d_1, d_2), thus the combination dose (d_1, d_2) is synergistic. Inversely, $f(d_1, d_2) > 0$ indicates that the combination dose (d_1, d_2) is antagonistic. Kong and Lee (6) used the different observed combination doses as the knots for the bivariate thin plate splines (12). The choice of knots works well when the number of combination doses is not large and the combination doses are not close, such as those from factorial designs or uniform designs (13). However, when ray designs are applied and the doses are low, the combination doses are very close and some columns of the design matrix (i.e., Ω and Z_1 in the following notations) may be highly correlated, which results in a nearly singular matrix for estimating the parameters in the function f. If that happens, a low rank smoothing thin plate spline (14) should be applied to avoid the singularity of the involved matrix due to the low rank of the design matrix. An example of such a low rank smoothing thin plate spline is the knots formed by selecting the observed combination doses with distances larger than some pre-specified small number. Alternatively, one may use an appropriate transformation of the dose, such as the log-transformation, to evenly distribute the experimental combination doses in certain regions in order to improve the ability to estimate the effect beyond additivity using bivariate thin plate splines.

Suppose the selected knots are $(\kappa_{1k}, \kappa_{2k})$ $(k=1,...,K)$, then the bivariate thin plate spline can be expressed by the following form:

$$f(d_1, d_2) = \gamma_0 + \gamma_1 d_1 + \gamma_2 d_2 + \sum_{k=1}^{K} v_k \eta\left(// (d_1, d_2)^T - (\kappa_{1k}, \kappa_{2k})^T //\right),$$

where $\gamma = (\gamma_0, \gamma_1, \gamma_2)^T$ and $v = (v_1, ..., v_K)^T$ are the parameters in the thin plate spline function f, and

$\eta(r) = \dfrac{1}{16\pi} r^2 \log r^2$ for $r > 0$ and $\eta(r) = 0$ for $r = 0$. The distance in the expression is the Euclidean distance. Let us denote

$$\Omega = \left[\eta\left(// (\kappa_{1k}, \kappa_{2k})^T - (\kappa_{1k'}, \kappa_{2k'})^T //\right)_{1 \leq k, k' \leq K}\right] \in R^{K \times K},$$

$$Y_R = \left[\left(E_1 - \hat{F}_p(d_{11}, d_{21})\right)1_{\{d_{11} \neq 0 \,\& \, d_{21} \neq 0\}}, ..., \left(E_n - \hat{F}_p(d_{1n}, d_{2n})\right)1_{\{d_{1n} \neq 0 \,\& \, d_{2n} \neq 0\}}\right]^T \in R^{n \times 1}$$

$$X = [1, d_{1i}, d_{2i}]_{1 \leq i \leq n} \in R^{n \times 3}, \quad Z_1 = \left[\eta\left(// (d_{1i}, d_{2i})^T - (\kappa_{1k}, \kappa_{2k})^T //\right)_{1 \leq k \leq K}\right]_{1 \leq i \leq n} \in R^{n \times K},$$

and $T^T = [1, \kappa_{1k}, \kappa_{2k}]_{1 \leq k \leq K} \in R^{K \times 3}$.

The coefficient γ and V can be obtained by minimizing the following penalized residual sum of squares:

$$(Y_R - X\gamma - Z_1 V)^T (Y_R - X\gamma - Z_1 V) + \lambda V^T \Omega V \quad (E6)$$

subject to $TV = 0$.

Following the notation by Kong and Lee (6) and Green and Silverman (12), consider a QR decomposition of T^T, say $T^T = FG$, where F is a $K \times K$ orthogonal matrix and G is a $K \times 3$ upper triangular matrix. Let F_1 be the first three columns of F, and F_2 be the last K-3 columns of F. We can show that $TV = 0$ if and only if V can be expressed as $F_2 \xi$ for some ξ. Thus, the penalized residual sum of squares can be expressed as

$$(Y_R - X\gamma - Z_1 F_2 \xi)^T (Y_R - X\gamma - Z_1 F_2 \xi) + \lambda \xi^T F_2^T \Omega F_2 \xi \cdot$$

Set $u = \left(F_2^T \Omega F_2\right)^{\frac{1}{2}} \xi$ and $Z = Z_1 F_2 \left(F_2^T \Omega F_2\right)^{-\frac{1}{2}}$, where $\left(F_2^T \Omega F_2\right)^{\frac{1}{2}}$ is the matrix square root of $F_2^T \Omega F_2$. The penalized residual sum of squares can be expressed as

$$(Y_R - X\gamma - Zu)^T (Y_R - X\gamma - Zu) + \lambda u^T u \quad (E7)$$

Based on the approach proposed by Ruppert, Wand, and Carroll (15) and Wang (16), and detailed by Kong and Lee (6) in this setting, the parameters in terms of γ and u can be obtained by solving the following mixed effect model:

$$Y_R = X\gamma + Zu + \varepsilon \quad (E8)$$

where

$$\begin{pmatrix} u \\ \varepsilon \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_u^2 I_{K-3} & 0 \\ 0 & \sigma_\varepsilon^2 I_n \end{pmatrix}\right) \cdot$$

Thus, the parameters can be estimated by $\begin{pmatrix} \hat{\gamma} \\ \hat{u} \end{pmatrix} = \left(C^T C + \hat{\lambda} D\right)^{-1} C^T Y_R$ with $\hat{\lambda} = \hat{\sigma}_\varepsilon^2 / \hat{\sigma}_u^2$, $C = [X \ Z]$, and $D = diag(0, 0, 0, 1, ..., 1)$, where the number of zeros in the matrix D corresponds to the number of $\gamma_i 's$ $(i=0,1,2)$ and the number of ones corresponds to the number of $u_i 's$ $(i = 1, ..., K - 3)$. Under these notations, for any combination dose (d_1, d_2), $f(d_1, d_2)$ can be predicted by $\hat{f}(d_1, d_2) = \hat{\gamma}_0 + \hat{\gamma}_1 d_1 + \hat{\gamma}_2 d_2 + Z_0 \hat{u}$ with

$$Z_0 = \left[\eta\left(// (d_1, d_2)^T - (\kappa_{1k}, \kappa_{2k})^T //\right)\right]_{1 \leq k \leq K} F_2 \left(F_2^T \Omega F_2\right)^{-1/2}, \text{ and}$$

an approximate $100(1-\alpha)\%$ point-wise confidence interval for $f(d_1, d_2)$ can be constructed by

$$\hat{f}(d_1, d_2) \pm z_{\alpha/2} \hat{\sigma}_\varepsilon \sqrt{C_d \left(C^T C + \hat{\lambda} D\right)^{-1} C_d^T}, \quad (E9)$$

where $C_d=(1, \ d_1, \ d_2, \ Z_0)$ and $z_{\alpha/2}$ is the upper $\alpha/2 \times 100\%$ percentile of the standard normal distribution. Thus, we can construct a 95% point-wise confidence band for $f = 0$ by taking the intercept lines of the lower and upper confidence surfaces with the dose plane. We then claim that the combination doses in the area outside the confidence bound with $f < 0$ are synergistic, the combination doses inside the bound are additive, and the combination doses in the area outside the bound with $f > 0$ are antagonistic.

Based on the 95% point-wise confidence intervals constructed from (*E9*), some combination doses that are additive may be claimed as synergistic or antagonistic. To be conservative and to control the family-wise error rate, we also construct the 95% lower and upper simultaneous confidence surfaces, which are based on a format similar to that of equation (*E9*) except that $z_{\alpha/2}$ is replaced by $\sqrt{EDF \times F^{\alpha}_{EDF, n-EDF}}$ (17), where *EDF* is the effective degrees of freedom from the resulting bivariate smoothing splines (12) and is defined as the trace of the matrix $C\left(C^T C + \hat{\lambda} D\right)^{-1} C^T$, and $F^{\alpha}_{EDF, n-EDF}$ is the upper $100 \times \alpha$ percentile of the *F* distribution with *EDF* and *n-EDF* degrees of freedom. Here *n* is the total number of observations except controls. A 95% simultaneous confidence band for $f = 0$ can be formed by taking the intercept lines of the 95% lower and upper simultaneous confidence surfaces with the dose plane. For the two case studies presented in the next section, we will present the plots of different patterns of drug interaction based on the respective 95% point-wise confidence intervals and the 95% simultaneous confidence band.

4. CASE STUDIES

The following two data sets were provided by Dr. William R. Greco. The data were collected during a study of the joint effect of trimetrexate (TMQ) and AG2034 on cells grown in medium with different concentrations of folic acid (FA): 2.3 μM in the first experiment (the low FA experiment), and 78 μM in the second experiment (the high FA experiment). TMQ is a lipophilic inhibitor of the enzyme dihydrofolate reductase, and AG2034 is an inhibitor of the enzyme glycinamide ribonucleotide formyltransferase. The unit of drug concentration is the micromole (μM) for all data analyzed in the case studies. The endpoint was the growth of human ileocecal adenocarcinoma (HCT-8) cells in 96-well assay plates as measured by the sulforhodamine B (SRB) protein stain. The drug treatments were randomly assigned to the cells in the assay wells. Each 96-well plate included 8 wells as instrumental blanks (no cells); thus 88 wells were used for drug treatments. Five replicate plates were used for each set of 88 treated wells. Each of the two large data sets were obtained from two 5-plate stacks with a maximum of 880 treated wells per experiment. Each experiment included 110 control wells, in which no drugs were applied to the cells. Ray designs were used, with the experimental doses

being distributed in 14 rays, including two rays for TMQ and AG2034 when used alone. The complete details and mechanistic implications of the study were reported by Faessel *et al* (18). Assuming that the first observation recorded in each dose or combination dose from the first 5-plate stack was from the same plate, say the 1st plate, the second observation from the 2nd plate, and so on, and also assuming that the first observation recorded in each dose or combination dose from the second 5-plate stack was from the same plate, say the 6th plate, the second observation from the 7th plate, and so on, we have a total of 10 plates for each of the two data sets.

To examine whether there is a significant difference among the plates, we applied one-way analysis of variance (ANOVA) to the controls in each data set. The p-values were 0.001 for the low FA experimental data and 0.005 for the high FA experimental data. The results indicate a significant plate effect among the 10 plates for each experiment, that is, the inter-plate variability is high. To attenuate the effect from the inter-plate variability, we applied a standardization procedure to each data set, dividing the effect readings by the mean of the controls in each associated plate. Thus, the mean for the controls within each plate was standardized to 1, and the effect for the controls was treated as 1. In addition to 110 controls for each experiment, the data included 761 observations for the low FA experiment and 769 observations for the high FA experiment. We applied the statistical method described in Section 3 to each of the two standardized data sets. We present the results for each experiment in the following two subsections.

Lee *et al* (19) performed extensive exploratory analyses on the same data sets and identified 129 outliers out of 871 (14.8%) effect readings in the low FA experiment and 126 outliers out of 879 (14.3%) effect readings in the high FA experiment. To compare our findings with the results previously obtained (19), we removed the outliers from the data and then again applied the statistical method described in Section 3. For each experiment, we report the detailed analyses of the original data set and of the modified data set that excluded the outliers.

4.1. Case study 1: cancer cells grown in a medium with 2.3 μM folic acid (low FA experiment)

In this experiment, the cells were grown in a medium with 2.3 μM folic acid. We fitted marginal dose-effect curves for TMQ and AG2034 by using both the median effect equation (*E3*) and the E_{max} model (*E5*). The dose levels for TMQ when applied alone were 5.47×10^{-6}, 4.38×10^{-5}, 1.38×10^{-4}, 4.38×10^{-4}, 8.75×10^{-4}, 1.75×10^{-3}, 3.5×10^{-3}, 7×10^{-3}, 2.21×10^{-2}, 7×10^{-2}, and 0.56 μM, and the dose levels for AG2034 when applied alone were 2.71×10^{-5}, 2.71×10^{-4}, 6.87×10^{-4}, 2.17×10^{-3}, 4.3×10^{-3}, 8.7×10^{-3}, 1.74×10^{-2}, 3.48×10^{-2}, 0.11, 0.3475, and 2.78 μM. Note that some effect readings at low doses or combination doses are greater than 1, thus, the logit transformation could not be carried out. We applied

Emax model and bivariate splines for drug interaction

Table 1. Estimated parameters for the E_{max} models in the case studies

Drug name	Low FA[1]			High FA[2]		
	E_{max}[3]	ED_{50}[4]	Slope m[5]	E_{max}[3]	ED_{50}[4]	Slope m[5]
TMQ[6]	0.8810 (0.0161)	0.0013 (0.0001)	2.2496 (0.2330)	0.8847 (0.0326)	0.0134 (0.0015)	3.7230 (0.7323)
AG2034[7]	0.8688 (0.0154)	0.0060 (0.0003)	3.1644 (0.3703)	0.8184 (0.0311)	0.4700 (0.0540)	1.6869 (0.2400)

[1]Estimated parameters for the marginal dose-effect curves in the low-concentration folic acid experiment; [2] estimated parameters for the marginal dose-effect curves in the high-concentration folic acid experiment; [3] maximum effect attributable to the drug; [4] dose level producing half of E_{max}; [5] factor (Hill coefficient) measuring sensitivity of the effect to the drug dose range; [6] trimetrexate; [7] experimental drug.

nonlinear least squares regression to estimate the parameters in models (*E3*) and (*E5*). Figures 3.A and 3.B show the respective fitted marginal dose-effect curves for TMQ and AG2034 with the dose levels shown on a log scale. The dotted-dashed lines are the curves based on the median effect model (*E3*), and the solid lines are the dose-effect curves based on the E_{max} model (*E5*). From the fitted dose-effect curves, we found that the E_{max} model (*E5*) provided a much better fit than the median effect equation for the marginal data. Therefore, we chose the E_{max} model (*E5*) to describe the dose-effect relationship in this case study. The parameters estimated for TMQ and AG2034 are shown in the three columns under the title "Low FA" in Table 1. Here, the estimate of $E_{max,TMQ}$ is slightly larger than the estimate of $E_{max,AG2034}$. We plotted the distribution of the combination doses using the original scale (not shown) and found that most of the combination doses were crowded in the region of the low doses, which could cause a singularity of the involved matrices due to the low rank of Ω and Z_1 used for estimating the effect beyond additivity when using bivariate thin plate splines (see Section 3.3). Hence, we applied a log transformation of the form log(dose+δ) for each dose level, where δ is a small number, say 2.74×10^{-6}, half of the smallest dose level for the two drugs when applied alone. We plotted the distribution of the combination doses on the log(dose+δ) scale, as shown in Figure 3.C. The points on the horizontal line in Figure 3.C are the doses of TMQ on the log(dose+δ) scale; the points on the vertical line are the doses of AG2034 on the log(dose+δ) scale; and the points on each of the remaining 12 design rays are the combination doses at each ray with each dose component on the log(dose+δ) scale. The 12 design rays for combination doses appearing left to right in Figure 3.C correspond to the combination doses at 12 ratios of TMQ to AG2034, i.e., 1:250, 1:125, 1:50, 1:20, 1:10, 1:5, 1:5, 2:5, 4:5, 2:1, 5:1, 10:1. The 12 rays are denoted by the letters E, F, G, H, I, J, K, L, M , N, O, P, which represent the respective curves 15, 13, 11, 7, 5, 3, 9, 4, 6, 10, 12, 14 in the original data set for the low FA experiment. Note that rays 3 and 9, denoted by J and K, indeed have the same fixed dose ratio. To obtain the predicted additive effects, we applied the procedure described in Section 3.2, keeping the dose levels on the original scale. The contour plot of the predicted additive effect is shown in Figure 3.D. Note that the effect levels for TMQ applied alone are obtained from (1-$E_{max,TMQ}$, 1), which is (0.1190, 1), and the effect levels for AG2034 applied alone are obtained from (1-$E_{max,AG2034}$, 1), which is (0.1312, 1). The vertical line with contour level 0.13 is the predicted effect produced by TMQ alone. The plot of the differences of the observed effects and predicted effects versus the dose levels of AG2034 on log(dose+δ)

scale is shown in Figure 3.E. That plot shows that the differences are not distributed around zero, rather, for some observations of AG2034, the differences are significantly less than zero, in the range of (-7, -4) on the log (dose+δ) scale with δ=2.74×10^{-6}, i.e., in the range of 0.001 μM to 0.018 μM on the original dose scale. Therefore, the pure additive effect model could not describe the data well. We used bivariate thin plate splines to fit the differences versus the transformed doses, with the knots at all the distinct transformed dose levels. The transformation is taken as log (dose+δ) for all single and combination doses. By convention, there is no drug interaction when a single drug is applied. Therefore the differences were set to zero for the marginal doses. Applying the bivariate thin plate splines (Section 3.3), we obtained $\hat{\sigma}_\varepsilon^2 = 0.0041$, $\hat{\sigma}_u^2 = 0.2318$, and $\hat{\lambda} = \hat{\sigma}_\varepsilon^2 / \hat{\sigma}_u^2 = 0.0178$. Next, we constructed 95% point-wise upper and lower confidence surfaces for the fitted bivariate spline function $f(d_1,d_2)$ based on equation (*E9*). Figure 3.F shows the contour plot of the fitted spline function $f(d_1,d_2)$ at the levels of -0.1, 0, and 0.1 as thin solid lines; the intercept lines of the corresponding 95% point-wise upper confidence surface with the dose plane as thick dashed lines; and the intercept lines of the corresponding 95% point-wise lower confidence surface with the dose plane as thick solid lines. The combination doses inside the thick dashed curves, shown in light blue, are synergistic because the effects beyond additivity at these combination doses are significantly smaller than zero. The combination doses inside the thick solid curves, shown in light pink, are antagonistic because the effects beyond additivity at these combinations were significantly larger than zero. The combination doses in the uncolored region, which lie between the thick solid curves and the thick dashed curves, are additive because the effects beyond additivity are not significantly different from zero. In particular, the combination doses with AG2034 in the transformed scale in the range of (-7, -4) inside the thick dashed lines are synergistic, which is consistent with the residual plot in Figure 3.E. The fitted response surface was obtained by adding the fitted spline function f (i.e., the effect beyond additivity) to the predicted additive surface. The contour plot of the fitted response surface at the contour levels of 0.2, 0.5, and 0.9 is shown in Figure 3.I. The final residuals were obtained by subtracting the fitted effects from the observed effects. The plots of final residuals versus the dose levels of TMQ and AG2034 on the log (dose+δ) scale are shown in Figures 3.G and 3.H, respectively. From these two plots, we see that the residuals are centered around zero along the experimental dose range. We conclude that the model fits the data reasonably well.
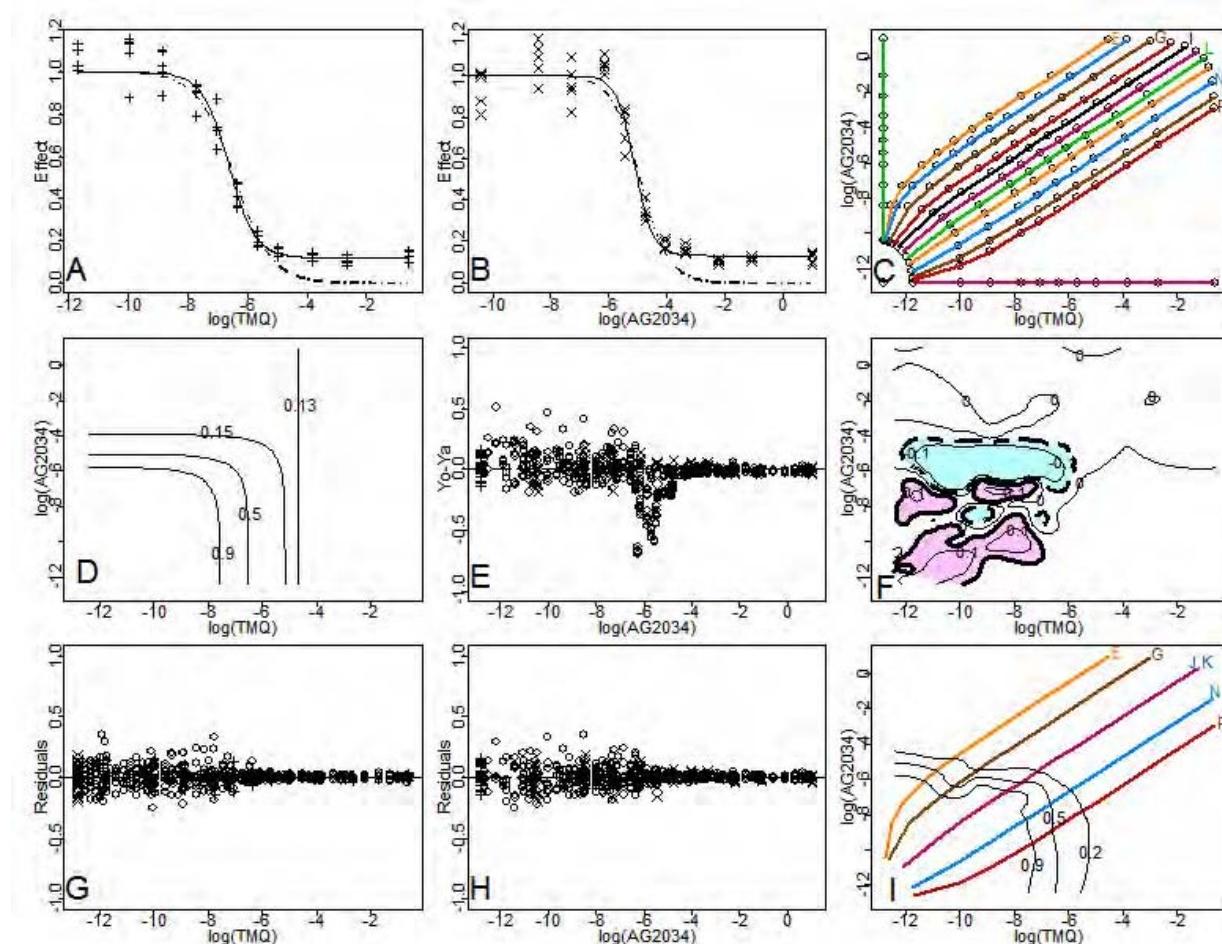
Figure 3. Analysis of the low FA experimental data. Panels A and B show the fitted marginal dose-effect curves for TMQ and AG2034 respectively, where the dotted-dashed line in each panel is the fitted dose-effect curve based on the median effect equation ($E\,3$), while the solid line in each panel is the fitted dose-effect curve based on the E_{max} model ($E\,5$). Panel C shows the distribution of the experimental doses and combination doses on the log (dose+δ) scale with $\delta = 2.74 \times 10^{-6}$, along with the 12 rays from left to right with dose ratios of TMQ versus AG2034 at 1:250, 1:125, 1:50, 1:20, 1:10, 1:5, 1:5, 2:5, 4:5, 2:1, 5:1, 10:1, denoted by the letters E, F, G, H, I, J, K, L, M , N, O, and P, representing the curves 15, 13, 11, 7, 5, 3, 9, 4, 6, 10, 12, and 14 in the original data set. Panel D shows the contour plot of the predicted additive effect, while Panel E shows the plot of the differences between the observed effects and the predicted effects versus the dose level of AG2034 on the log (dose+δ) scale. Panel F shows the contour plot of the fitted effect beyond the additivity effect at levels -0.1, 0, and 0.1 as thin solid lines, along with the intercept line of the 95% point-wise upper confidence surface with the dose plane as thick dashed lines and the intercept line of the 95% point-wise lower confidence surface with the dose plane as thick solid lines. In Panel F, the combination doses in the light blue area are synergistic, the combination doses in the light pink area are antagonistic, and the combination doses in the uncolored area are additive. The colored lines in Panels C and I are the design rays. Panels G and H are the plots of the final residuals versus TMQ and AG2034 on the log (dose+δ) scale, respectively, and Panel I is the contour plot of the fitted response surface at the levels of 0.9, 0.5, and 0.2, along with some representative design rays.

To examine the patterns of drug interactions in different rays and different experimental combination doses, we combined Figures 3.F and 3.I, that is, we plotted the contour curves of the fitted response surface at the levels of 0.2, 0.5, and 0.9 in Panel F. We also plotted the representative design rays, with the experimental combination doses shown as dots on these rays (Figure 4.A), As seen in that figure, the combination doses on the rays E through K (curves 15, 13, 11, 7, 5, 3, 9 in the original data set) are synergistic when the effect levels are between 0.9 and a number smaller than 0.2. The

combination doses on these rays are additive when the effect level is less than this small number, and the combination doses at low levels on these lines are either additive or antagonistic. The combination doses on the rays N, O, and P (curves 10, 12, and 14 in the original data set) are additive when the effects are less than 0.9, and the combination doses at low dose levels are antagonistic.

In addition to the 95% point-wise confidence interval, we constructed the 95% simultaneous confidence
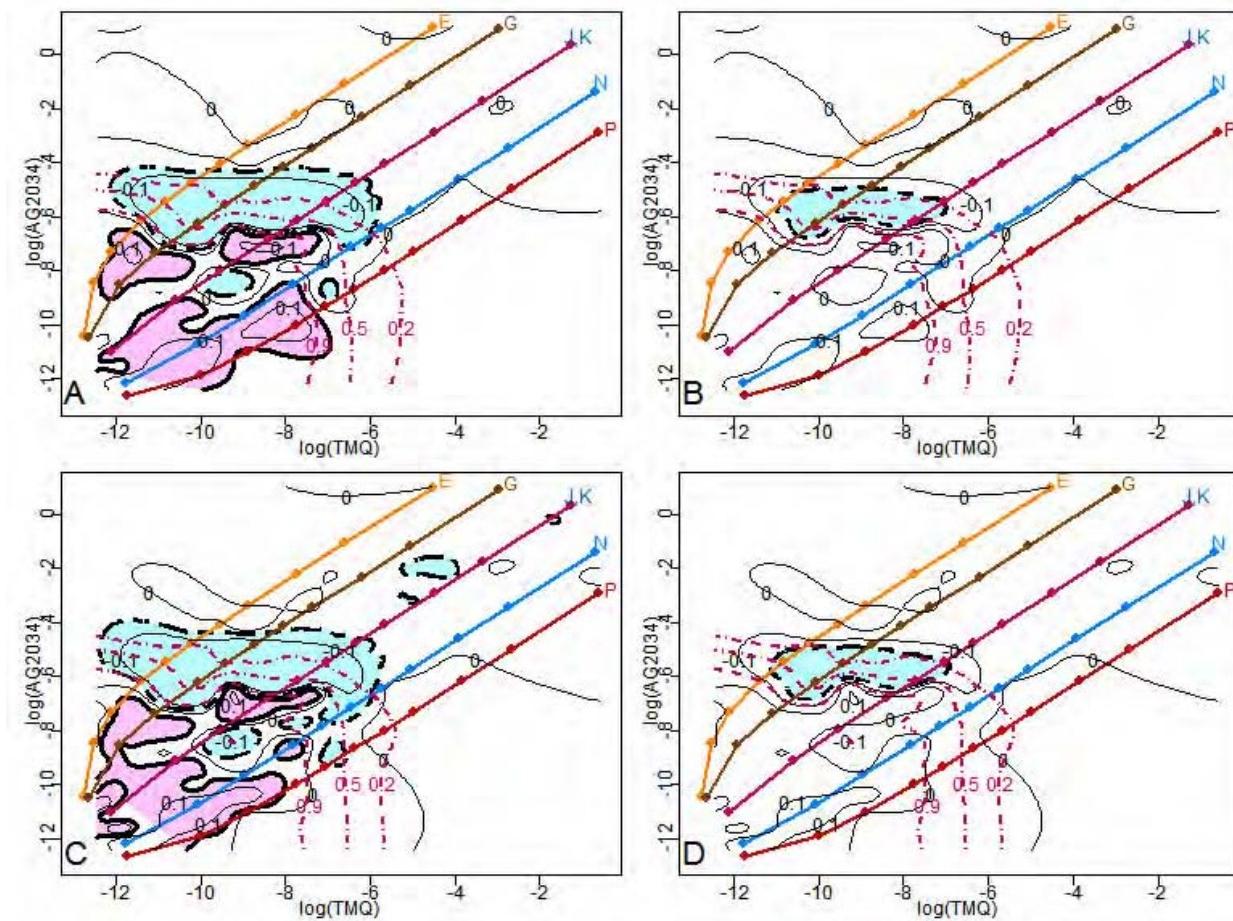
Figure 4. Different patterns of drug interactions for the low FA experimental data. Panel A is based on 95% point-wise confidence intervals; Panel B is based on 95% simultaneous confidence band. Panel A is the combination of Figures 3.F and 3.I, along with the design points shown as dots on each ray. Thin solid lines are contour lines of the fitted effect surface beyond the additivity surface at the levels of -0.1, 0, and 0.1; thick dashed lines are the intercept lines of the upper 95% point-wise confidence surface with the dose plane; thick solid lines are the intercept lines of the lower 95% point-wise confidence surface with the dose plane. The colored lines labeled E, G, J, K, N, and P are the representatives of the design rays. The red dotted-dashed lines are the contour lines of the fitted response surface at the levels of 0.9, 0.5, and 0.2, respectively. Based on Panel A, the combination doses in the light blue area are synergistic, the combination doses in the light pink area are antagonistic, and the combination doses on the uncolored area are additive. Panel B presents the same information as Panel A except that the thick dashed lines are the intercept lines of the upper 95% simultaneous confidence surface with the dose plane and there are no intercept lines for the lower 95% simultaneous confidence surface with the dose plane. Based on Panel B, the combination doses inside the dashed lines are synergistic, otherwise additive. Panel B gives more conservative results for assessing drug interactions. Panels C and D are the results from fitting the data set excluding outliers for the low FA experiment, where the information in Panel C is parallel to that in Panel A, and the information in Panel D is parallel to that in Panel B.

band based on equation (E9) with $z_{\alpha/2}$ replaced by $\sqrt{EDF \times F^{\alpha}_{EDF,n-EDF}}$. Here EDF=119, n=761, $\sqrt{EDF \times F^{\alpha}_{EDF,n-EDF}}$ = 12.20, and $\alpha = 0.05$. The resulting patterns of drug interactions are shown in Figure 4B, in which the thick dashed line is the intercept line of the 95% upper simultaneous confidence surface with the dose plane. Based on Figure 4B, we conclude that the combination doses inside the thick dashed curves, shown in light blue, are synergistic. The combination doses outside the thick dashed curves are

additive. As seen in the figure, compared to the point-wise confidence interval approach, the simultaneous confidence band method shrinks the synergistic area and results in the disappearance of the antagonistic area. A point-wise confidence interval is appropriate for making inferences for each observed combination dose. The simultaneous confidence band is suitable for making a global assessment; however, it can be overly conservative.

Following those analyses, we evaluated the low FA experimental data set from which we had removed the

outliers (19). The results of our assessment of drug interactions for this data set are presented in Figures 4.C and 4.D. Figures 4.A and 4.C contain parallel information, as do Figures 4.B and 4.D. Comparing the plots across the panels, we conclude that the results from fitting the original data set and those from fitting the data set excluding outliers are very similar. Therefore, the semiparametric method presented in Section 3 is robust to outliers in this example.

We recommend using caution when considering extrapolations based on spline estimations. The fitted response surface for the differences between the observed effects and predicted effects gives an overall picture of the drug interactions (see Figures 4.A and 4.B). However, the fitted results on the two larger areas outside experiment rays E and P should not be over-interpreted because (i) there are no experimental data in those areas and (ii) we forced the differences of the observed effects and predicted additive effects to be zero at the marginal observed dose levels.

4.2. Case study 2: cancer cells grown in a medium with 78 μM folic acid (high FA experiment)

In the high FA experiment, the dose levels for TMQ when applied alone were 5.47×10^{-6}, 4.38×10^{-5}, 1.38×10^{-4}, 4.38×10^{-4}, 8.75×10^{-4}, 1.75×10^{-3}, 3.5×10^{-3}, 7×10^{-3}, 2.21×10^{-2}, 7×10^{-2}, and 0.56 μM, and the dose levels for AG2034 when applied alone were 2.71×10^{-4}, 2.17×10^{-3}, 6.87×10^{-3}, 2.17×10^{-2}, 4.34×10^{-2}, 8.68×10^{-2}, 1.74×10^{-1}, 3.47×10^{-1}, 1.1, 3.47, and 27.8 μM. The procedure we used to analyze this data set was the same as that used in case study 1. By applying nonlinear least squares regression, we estimated the marginal dose-effect curves using the median effect equation (*E3*) (dotted-dashed lines) and the E_{max} model (*E5*) (solid lines), (shown in Figures 5.A and 5.B). It is clear that the E_{max} model (*E5*) fits the data better than the median effect equation, thus, we chose the E_{max} model (*E5*) as the dose-effect curve for this data set. The estimated parameters for the marginal dose-effect curves for the E_{max} model are shown in the three columns under the title "High FA" in Table 1. The combination doses on the original scale (not shown) are crowded in the region of the low doses, thus we applied the transformation in the form of log(dose+δ) to each dose level, where δ is a small number, say 2.74×10^{-6}, one half of the lowest dose level for TMQ and AG2034 when applied alone. The distribution of the experimental dose levels on the log(dose+δ) scale is shown in Figure 5.C. The 12 design rays for the combination doses correspond to the 12 dose ratios of TMQ versus AG2034 at 1:2500, 1:1250, 1:500, 1:200, 1:100, 1:50, 1:50, 1:25, 1:12.5, 1:5, 1:2, and 1:1, which are denoted by the letters E, F, G, H, I, J, K, L, M , N, O, and P, representing the curves 15, 13, 11, 7, 5, 3, 9, 4, 6, 10, 12, 14 in the original data set for the high FA experiment. Applying the procedure described in Section 3.2, we obtained the contour plot of the predicted additive effect that is shown in Figure 5.D. We see that the contour line at level 0.15 is the predicted effect produced by TMQ alone because AG2034 could not produce such an effect when applied alone; the effect levels for AG2034 applied alone range from 0.1816 to 1. Figure 5.E shows the

differences of the observed effects and predicted effects versus the dose levels of AG2034 on the log(dose+δ) scale. That plot shows that the differences are not centered around zero, rather for some observations of AG2034, the differences are significantly less than zero, in the range of (-5, 0) on the log(dose+δ) scale, i.e., in the range of 6.7×10^{-3} μM to 1.0 μM on the original dose scale. These findings indicate that some combination doses were synergistic and that the pure additive effect model could not describe the data well. We used bivariate thin plate splines to fit these differences versus the transformed doses or combination dose, with the knots at all the distinct transformed dose levels. The transformation is taken as log (dose+δ) for all single doses and combination doses. We constructed its 95% point-wise confidence surfaces based on equation (*E9*). The estimated $\hat{\sigma}_{\varepsilon}^2 = 0.0066$, $\hat{\sigma}_u^2 = 0.0779$, and $\hat{\lambda} = \hat{\sigma}_{\varepsilon}^2 / \hat{\sigma}_u^2 = 0.0842$. Figure 5.F shows the contour plot of the fitted spline function f at the levels of -0.1, 0, and 0.1 as thin solid lines; the intercept lines of its corresponding 95% point-wise upper confidence surface with the dose plane as thick dashed lines; and the intercept lines of its corresponding 95% point-wise lower confidence surface with the dose plane as thick solid lines. The combination doses inside the thick dashed curves, shown in light blue, are synergistic; the combination doses inside the thick solid curves, shown in light pink, are antagonistic; and the combination doses in the uncolored area are additive. We obtained the fitted response surface by adding the fitted spline function f to the predicted additive surface (shown in Figure 5.I). The plots of the final residuals versus the dose levels of TMQ and AG2034 on the log(dose+δ) scale are shown in Figures 5.G and 5.H, respectively. From these two plots, we see that the residuals are centered around zero along the experimental dose range, indicating that the model describes the data reasonably well.

To examine the patterns of drug interactions in different rays and different experimental combination doses, we combined the plots in Figures 5.F and 5.I to form Figure 6.A (as we did when analyzing the data from the low FA experiment). From Figure 6.A, we see that the combination doses on all 12 rays are synergistic when the effect levels are between 0.9 and 0.15. The combination doses at high dose levels are additive, and most of the combination doses at low dose levels are additive. In addition, we constructed a 95% simultaneous confidence band based on equation (*E9*) with $z_{\alpha/2}$ replaced by $\sqrt{EDF \times F_{EDF, n-EDF}^{\alpha}}$. Here $EDF=91$, n=769, and $\sqrt{EDF \times F_{EDF, n-EDF}^{\alpha}} = 10.77$. The results are presented in Figure 6.B, in which the thick dashed line is the intercept line of the upper 95% simultaneous confidence surface with the dose plane. Based on Figure 6.B, we conclude that the combination doses inside the thick dashed curves, shown in light blue, are synergistic. The combination doses outside the thick dashed curves are additive. As in our analysis of the data from case study 1, in this analysis, we found the simultaneous confidence band to yield more conservative results and to be more suitable for a global assessment. For this case study, we also assessed the data set from which we had removed the outliers. The results for assessing drug
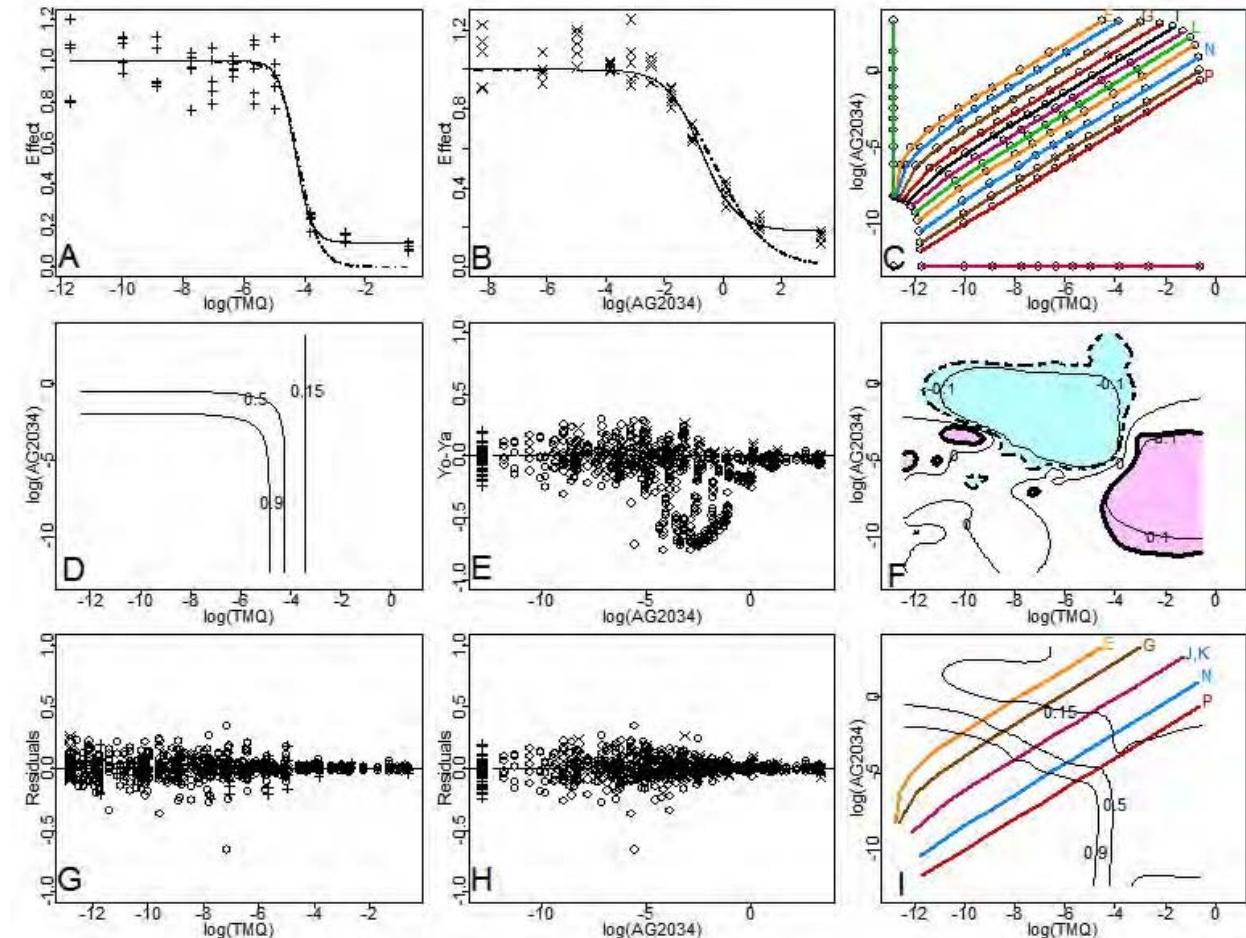
Figure 5. Analysis of the high FA experimental data. Panels A and B show the fitted marginal dose-effect curves for TMQ and AG2034 respectively; the dotted-dashed line is the fitted dose-effect curve based on the median effect equation ($E\,3$); the solid line is the fitted dose-effect curve based on the E_{max} model ($E\,5$). Panel C shows the distribution of the experimental doses and combination doses on the log(dose+δ) scale with $\delta = 2.74 \times 10^{-6}$, along with the 12 rays (left to right) with dose ratios of TMQ versus AG2034 at 1:2500, 1:1250, 1:500, 1:200, 1:100, 1:50, 1:50, 1:25, 1:12.5, 1:5, 1:2, 1:1, denoted by the letters E, F, G, H, I, J, K, L, M , N, O, P, representing the curves 15, 13, 11, 7, 5, 3, 9, 4, 6, 10, 12, 14 in the original data set. Panel D shows the contour plot of the predicted additive effect. Panel E shows the plot of the differences between the observed effects and the predicted effects versus the dose level of AG2034 on the log(dose+δ) scale. Panel F shows the contour plot of the fitted effect beyond the additivity effect at levels -0.1, 0, and 0.1, along with the intercept line of the upper 95% point-wise confidence surface with the dose plane as thick dashed lines and the intercept lines of the lower 95% point-wise confidence surface with the dose plane as thick solid lines. In Panel F, synergistic combination doses are in light blue; antagonistic combination doses are in light pink; additive combination doses are in the uncolored area. The colored lines in Panels C and I represent the design rays. Panels G and H are the plots of the final residuals versus TMQ and AG2034 on the log(dose+δ) scale, respectively. Panel I is the contour plot of the fitted response surface at the levels of 0.9, 0.5, and 0.15, along with some representative design rays.

interactions are presented in Figures 6.C and 6.D. Comparing Figures 6.A to C and 6.B to D, we conclude that the results from fitting the original data set and those from fitting the data set excluding outliers are very similar. Thus, the results indicate that the semiparametric method is robust to outliers.

5. SUMMARY AND PERSPECTIVE

We extended the approach proposed by Kong and Lee (6) to a situation for which the E_{max} model is more appropriate to describe the marginal dose-effect relationship. We considered the possibility that some effect readings at low doses may fall beyond the mean of the controls. Under such circumstances, the standardized effect is greater than 1 and a logit transformation to a linear model (4, 8, 9) cannot be carried out. Hence, other models such as the E_{max} model are needed and nonlinear least squares regression methods can be applied for estimating parameters for the dose-effect curves. We applied nonlinear least squares regression in the case studies to estimate the parameters for the dose-effect curves specified by the median effect equation and the E_{max} model.
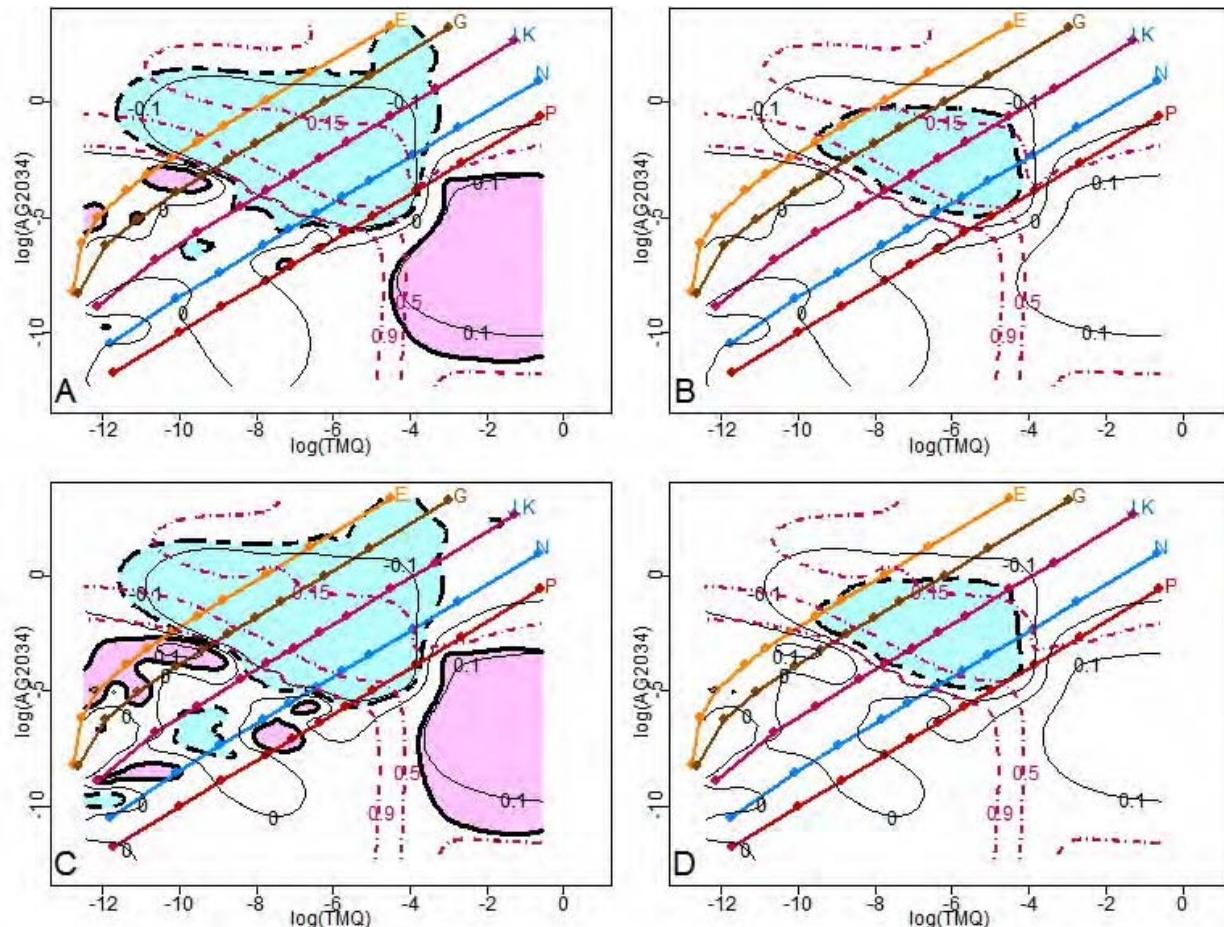
Figure 6. Different patterns of drug interactions for the high FA experiment. Panel A is based on 95% point-wise confidence intervals; Panel B is based on 95% simultaneous confidence band. Panel A is the combination of Figures 5.F and I, along with the design points shown as dots on each ray. Thin solid lines are contour lines of the fitted effect surface beyond the additivity surface at the levels of -0.1, 0, and 0.1; thick dashed lines are the intercept lines of the upper 95% point-wise confidence surface with the dose plane; thick solid lines are the intercept lines of the lower 95% point-wise confidence surface with the dose plane. The colored lines labeled E, G, J, K, N, and P represent the design rays; red dotted-dashed lines are the contour lines of the fitted response surface at the levels of 0.9, 0.5, and 0.15. In Panel A, the synergistic combination doses are in light blue; the antagonistic combination doses are in light pink; additive combination doses in the uncolored area. Panel B gives the same information as Panel A except that the thick dashed lines are the intercept lines of the upper 95% simultaneous confidence surface with the dose plane. Based on Panel B, the combination doses inside the dashed lines are synergistic, otherwise additive. Panel B gives more conservative results for assessing drug interactions. Panels C and D are the results excluding outliers, and are parallel to the results in Panels A and B, respectively.

Additionally, we extended the approach of Kong and Lee (6) as a solution to the problem arising when the experimental points are very close and the low rank of the design matrix may cause computational problems in matrix inversion. For this situation, we considered a low-rank thin plate spline (14) to estimate the surface beyond additivity, and, alternatively, we applied an appropriate transformation to the doses so that the combination doses on the transformed scale were more evenly distributed. In the case studies, we first applied the transformation $\log(dose+\delta)$ to each component of the combination doses and then applied bivariate thin plate splines with knots representing all the different observed doses on the $\log(dose+\delta)$ scale. In both

case studies, we chose δ as half of the smallest non-zero dose among TMQ and AG2034 when applied alone, that is, $\delta=2.74\times10^{-6}$ for both experiments. The value of δ selected should not be too small or too large compared with the magnitude of the dose levels. An extremely small δ results in a relatively large distance between the marginal doses and combination doses. Conversely, a large δ dominates in the transformation $\log(dose+\delta)$ when the dose levels are low. From the final residual plots, it is evident that the current transformation works well.

It is well known that the smoothing parameter λ governs the trade-off between the goodness-of-fit and the smoothness of the function f. When λ becomes larger, the

fitted function f tends to be smoother and the residuals tend to be larger. The selection of the smoothing parameter plays a key role in the fitted results. In the case studies, the smoothing parameter, $\hat{\lambda}$, was selected as $\hat{\sigma}_\varepsilon^2 / \hat{\sigma}_u^2$, which is almost identical to the selected smoothing parameter based on the generalized cross validation criterion and "leave-out-one" cross validation criterion. For example, for the low FA experimental data, the selected parameters based on the mixed model approach used in this paper, cross validation criterion, and generalized cross validation criterion were 0.0178, 0.0112, and 0.0071, respectively. For the high FA experimental data, the corresponding selected parameters were 0.0842, 0.0842, and 0.0531, respectively. Indeed, Kohn, Ansley, and Tharm (20) showed that the estimation of the smoothing parameter based on a mixed model approach is comparable with the standard method of the generalized cross validation criterion. By applying a mixed effects model, the smoothing parameter can be automatically determined by $\hat{\sigma}_\varepsilon^2 / \hat{\sigma}_u^2$. This method has been implemented in S-PLUS by Ruppert *et al* (15) using the lme function (21). In our previous study (1), based on extensive simulations, we showed that the selection of the smoothing parameter provides a good estimate to the underlying function in general.

In the two case studies, we also performed the same analyses for the two reduced data sets analyzed by Lee *et al* (19), and achieved almost identical results, which indicates that this semiparametric method is robust to outliers. The semiparametric method we have developed can also be used to assess drug interactions for the combination doses not on the design rays, and to identify complex patterns of drug interaction in combination studies. In addition, the method gives an overall assessment of the combination effect in the entire two-dimensional dose space spanned by the experimental doses with a caveat that extrapolation beyond data points can be risky.

We also note that the estimated function $f(d_1,d_2)$ and its 95% confidence surfaces can guide the exploration of whether some parametric models are sufficient to describe the data. Many parametric models have been proposed in the literature. Greco, Bravo, and Parsons (3) provided an excellent review of the response surface approach. However, without prior knowledge of the response surface model, or adequate representation of the data by a parametric model, most parametric approaches will fail. Blindly using any parametric model can be dangerous and may lead to wrong conclusions of drug interactions. In our proposed approach, there is no need to assume any parametric models for $f(d_1,d_2)$. We provide a promising approach by modeling the mixture effect data with spline techniques via a mixed-effect model. We advocate the use of the semiparametric method for model building because the true patterns of drug interactions are typically not known. The conclusions regarding drug interactions are based on the estimated f and its confidence surfaces, which are determined by the underlying data. The S-PLUS code used to evaluate the case studies can be obtained from the first author.

7. REFERENCES

1. Fumihiko Kanzawa, Kazuto Nishio, Kazuya Fukuoka, Toshihiko Sunami, and Nagahiro Saijo: *In vitro* interactions of a new derivative of spicamycin, KRN5500, and other anticancer drugs using a three-dimensional model. *Cancer Chemother Pharmacol* 43, 353-363 (1999)

2. Morris C. Berenbaum: What is synergy? *Pharmacol Rev* 41, 93-141 (1989)

3. William R. Greco, Gregory Bravo, John C. Parsons: The search of synergy: A critical review from a response surface perspective. *Pharmacol Rev* 47(2), 331-385 (1995)

4. Ting-Chao Chou: Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol Rev* 58, 621-681 (2006)

5. J. Jack Lee, Maiying Kong, Gregory D. Ayers, Reuben Lotan: Interaction index and different methods for determining drug interaction in Combination Therapy. *J Biopharm Stat* 17, 461-480 (2007)

6. Maiying Kong, J. Jack Lee: A semiparametric response surface model for assessing drug interactions. *Biometrics* 64, 396-405 (2008)

7. Maiying Kong, J. Jack Lee: A general response surface model with varying relative potency for assessing drug interactions. *Biometrics* 62 (4), 986-995 (2006)

8. J. Jack Lee, Maiying Kong: A confidence interval for interaction index for assessing multiple drug interaction. *Stat Biopharm Res* 1, 4-17 (2009)

9. Ting-Chao Chou, Paul Talalay: Quantitative analysis of dose effect relationships: the combined effects of multiple drugs or enzyme inhibitors. *Adv Enzyme Regul* 22, 27-55 (1984)

10. Naitee Ting: Dose Finding in Drug Development. Springer, New York, USA, 127-145 (2006)

11. Daniël M. Jonker, Sandra A. G. Visser, Piet H. van der Graaf, Rob A. Voskuyl, and Meindert Danhof:

Towards a mechanism based analysis of pharmacodynamic drug-drug interaction *in vivo*. *Pharmacol and Ther* 106, 1-18 (2005)

12. Peter J. Green, Bernard W. Silverman: Nonparametric Regression and Generalized Linear Models. Chapman & Hall, London (1994)

13. Ming Tan, Hong-Bin Fang, Guo-Liang Tian, and Peter J. Houghton: Experimental design and sample size determination for testing synergism in drug combination studies based on uniform measures. *Stat Med* 22, 2091-2100 (2003)

14. Haonan Wang, and M. Giovanna Ranalli: Low-rank smoothing splines on complicated domains. *Biometrics* 63, 209-217 (2007)

15. David Ruppert, Matthew P. Wand, and Raymond J. Carroll: Semiparametric Regression. Cambridge University Press, UK (2003)

16. Yuedong Wang: Mixed effect smoothing spline analysis of variance. *J Roy Stat Soc B* 60, 159-174 (1998).

17. Henry Scheffe: The Analysis of Variance. New York: John Wiley & Sons (1959)

18. Hélène M. Faessel, Harry K. Slocum, Robert C. Jackson, Theodore J. Boritzki, Youcef M. Rustum, M. G. Nair, and William R. Greco: Super *in vitro* synergy between inhibitors of dihydrofolate reductase and inhibitors of other folate-requiring enzymes: The critical role of polyglutamylation. *Cancer Res* 58, 3036-3050 (1998)

19. J. Jack Lee, Heather Y. Lin, Diane D. Liu, and Maiying Kong: Applying E_{max} model and interaction index for assessing drug interaction in combination studies. *Front Biosci* (2009) [in press]

20. Robert Kohn, Craig F. Ansley, David Tharm: The performance of cross validation and maximum likelihood estimators of spline smoothing parameters. *J Am Stat Assoc* 86, 1042-1050 (1991)

21. José C. Pinheiro, Douglas M. Bates: Mixed-Effects Models in S and S-PLUS. New York: Springer-Verlag (2000)

Abbreviations: ANOVA = analysis of varianceFA = folic acidHCT-8 cells = human illeocecal adenocarcinoma cellslme = linear mixed effectsμM = micromoleTMQ = trimetrexateSRB = sulforhodamine B

Key Words: Additivity, Antagonism, Bivariate splines, Combination therapy, E_{max} model, the Loewe additivity model, Synergy, Review

Send correspondence to: Maiying Kong, Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville, Louisville, Kentucky 40292, U.S.A. Tel: 502-852-3988, Fax: 502-852-3294, E-mail: maiying.kong@louisville.edu

E_{max} **model and interaction index for assessing drug interaction in combination studies**

J. Jack Lee,[1] Heather Y. Lin,[1] Diane D. Liu,[1], Maiying Kong[2]

[1]*Department of Biostatistics, The University of Texas M. D. Anderson Cancer Center, Unit 1411, 1515 Holcombe Boulevard, Houston, Texas 77030, U.S.A.,* [2]*Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville, Louisville, Kentucky 40292, U.S.A.*

TABLE OF CONTENTS

1. ABSTRACT

Applying the E_{max} model in a Lowe additivity model context, we analyze data from a combination study of trimetrexate (TMQ) and AG2034 (AG) in media of low and high concentrations of folic acid (FA). The E_{max} model provides a sufficient fit to the data. TMQ is more potent than AG in both low and high FA media. At low TMQ:AG ratios, when a smaller amount of the more potent drug (TMQ) is added to a larger amount of the less potent drug (AG), synergy results. When the TMQ:AG ratio reaches 0.4 or larger in low FA medium, or when the TMQ:AG ratio reaches 1 or larger in high FA medium, synergy is weakened and drug interaction becomes additive. In general, synergistic effect in a dilution series is stronger at higher doses that produce stronger effects (closer to $1-E_{max}$) than at lower dose levels that produce weaker effects (closer to 1). The two drugs are more potent in the low compared to the high FA medium. Drug synergy, however, is stronger in the high FA medium.

2. INTRODUCTION

Due to complex disease pathways, combination treatments can be more effective and less toxic than treatments with a single drug regimen. Successful applications of combination therapy have improved treatment effectiveness for many diseases. For example, the combination of a non-nucleoside reverse transcriptase inhibitor or protease inhibitor with two nucleosides is considered a standard front-line therapy in the treatment of AIDS. Typically, a combination of three to four drugs is required to provide a durable response and reconstitution of the immune system (1). Another example is platinum-based doublet chemotherapy regimens as the standard of care for patients with advanced stage non–small-cell lung cancer (2). Combination treatments have also been shown to prevent and to overcome drug resistance in infectious diseases such as malaria and in complex diseases such as cancer (3, 4). Emerging developments in cancer therapy involve combining multiple targeted agents with or without

chemotherapy, or combining multiple treatment modalities such as drugs, surgical procedures, and/or radiation therapy (5, 6).

How do we assess the effect of a combination therapy? It is a simple question, yet it requires a complex answer. A superficial way to answer the question is to determine that a combination therapy is working if its effect is greater than that produced by each single component given alone. The notion of classifying drug interaction as additive, synergistic, or antagonistic is logical and easily understood in a general sense, but can be confusing in specific application without consensus on a standard definition. Excellent reviews of drug synergism have been written by Berenbaum (7), Greco *et al* (8), Suhnel (9), Chou (10), and Tallarida (11), to name a few. In essence, to quantify the effect of combination therapy, we must first define drug synergy in terms of "additivity." An effect produced by a combination of agents that is more (or less) than the additive effect of the single agents is considered synergistic (or antagonistic). Then, we must further assess drug interaction in a statistical sense. Under a more rigorous definition, synergy occurs when the combined drug effect is statistically significantly higher than the additive effect. Conversely, antagonism occurs when the combination effect is statistically significantly lower than the additive effect.

Despite the controversy arising from multiple definitions of additivity or no drug interaction, the Loewe additivity model is commonly accepted as the gold standard for quantifying drug interaction (7-11). The Loewe additivity model is defined as

$$\frac{d_1}{D_{y,1}} + \frac{d_2}{D_{y,2}} = 1 \; . \qquad (E1)$$

Here y is the predicted additive effect at the combination dose (d_1, d_2) when the two drugs do not interact. $D_{y,1}$ and $D_{y,2}$ are the respective doses of drug 1 and drug 2 required to produce the same effect y when used alone. Note that the Loewe additivity can be easily demonstrated in a "sham combination" (i.e., a drug combined with itself or its diluted form). For example, suppose drug 2 is a 50% diluted form of drug 1. The combination of one unit of drug 1 and one unit of drug 2 will produce the same effect as 1.5 units of drug 1 or 3 units of drug 2. Plugging the respective values in equation $(E1)$, we have $1/1.5 + 1/3 = 1$. Given the dose-effect relationship for each single agent, say $E_i(d) = f_i(d)$ for agent i $(i=1,2)$, $D_{y,i}$ can be obtained by using the inverse function of f_i, say, $f_i^{-1}(y)$. Replacing $D_{y,1}$ and $D_{y,2}$ in equation $(E1)$ with $f_1^{-1}(y)$ and $f_2^{-1}(y)$, respectively, we can rewrite equation $(E1)$ as

$$\frac{d_1}{f_1^{-1}(y)} + \frac{d_2}{f_2^{-1}(y)} = 1 \; . \qquad (E2)$$

Note that $(E2)$ involves an unknown variable y. By solving equation $(E2)$, the predicted additive effect y_{add} can be obtained under the Loewe additivity model. Denote that the observed mean effect is y_{obs} at the combination dose (d_1, d_2). The drug combination at that dose is considered synergistic, additive, or antagonistic when the effect y_{obs} is

greater than, equal to, or less than y_{add}, respectively. When the dose-effect curve is decreasing (or increasing), a synergistic effect corresponds to a smaller (or larger) value than the predicted quantity.

Alternatively, to measure and quantify the magnitude of drug interaction, the interaction index (II) can be defined as

$$II = \frac{d_1}{D_{y_{obs},1}} + \frac{d_2}{D_{y_{obs},2}} \qquad (E3)$$

Note that $II < 1$, $II = 1$, and $II > 1$ correspond to the drug interaction being synergistic, additive, and antagonistic, respectively. Chou and Talalay (12) proposed the following median effect equation $(E4)$ to characterize the dose-effect relationship in combination studies:

$$E(d) = \frac{(d/ED_{50})^m}{1 + (d/ED_{50})^m} \; , \qquad (E4)$$

where ED_{50} is the dose required to produce 50% of the maximum effect. Although the median effect equation can be applied in many settings, it assumes that when m is positive, $E(d)=0$ for $d=0$ and $E(d)=1$ for $d=\infty$. On the other hand, when m is negative, $E(d)=1$ for $d=0$ and $E(d)=0$ for $d=\infty$. If we assume that the data follow the median effect equation, a linear relationship can be found by plotting the logit transformation of the effect versus the logarithm transformed dose. A more detailed account of the interpretation and use of the interaction index can be found in a number of references (13-16). Several methods for constructing the confidence interval estimation of the interaction index were proposed by Lee and Kong (17).

To help advance research developing and comparing methods for analyzing data for combination studies, Dr. William R. Greco at the Roswell Park Cancer Institute has organized an effort and invited several groups to participate in an exercise to compare rival modern approaches to model data from two-agent concentration-effect studies. We describe the data and statistical methods, including the E_{max} model, and the calculation of the interaction index under the E_{max} model in Section 3. We describe an exploratory data analysis in Section 4, and data preprocessing for outlier rejection and standardization in Section 5. We present the main results of the data analysis in Section 6 and summarize our findings in Section 7. We close with a discussion in Section 8.

3. STATISTICAL METHODS

3.1. Data sets

Two data sets provided by Dr. Greco are used to examine the effect of the combination treatment of trimetrexate (TMQ) and AG2034 (AG) in HCT-8 human ileocecal adenocarcinoma cells. The cells were grown in a medium with two concentrations of folic acid: 2.3 μM (the first data set, called low FA) and 78 μM (the second data set, called high FA). Trimetrexate is a lipophilic inhibitor

of the enzyme dihydrofolate reductase; and AG2034 is an inhibitor of the enzyme glycinamide ribonucleotide formyltransferase. The experiment was conducted on 96-well plates. The endpoint was cell growth measured by an absorbance value (ranging from 0 to 2) and recorded in an automated 96-well plate reader. Each 96-well plate included 8 wells as instrumental blanks (no cells); the remaining 88 wells received drug applications. The experiments were performed using the "ray design," which maintains a fixed dose ratio between TMQ and AG in a series of 11 dose dilutions. With 88 wells in each plate, each 5-plate stack allowed for an assessment of the combination doses at 7 curves (i.e., design rays) plus a "curve" with all controls. Two stacks were used for studying 14 design rays: TMQ only, AG only, and twelve other design rays with a fixed dose ratio (TMQ:AG) for each ray. The fixed dose ratios in the low FA experiment were 1:250, 1:125, 1:50, 1:20, 1:10, 1:5 (2 sets), 2:5, 4:5, 2:1, 5:1, and 10:1. The fixed dose ratios in the high FA experiment were 1:2500, 1:1250, 1:500, 1:200, 1:100, 1:50 (2 sets), 1:25, 2:25, 1:5, 1:2, and 1:1. Data from each of the 16 curves (2 for controls, 2 for single agents, and 12 for combinations) are grouped together. Curves 1-8 were performed on the first stack with curve 8 serving as the "control" experiment while curves 9-16 were performed on the second stack with curve 16 serving as the "control" experiment. The assignment of different drug combinations to the cells in the wells was randomized across the plates. Five replicate plates were used for each set of two stacks, resulting in a total of 10 plates for each of the two medium conditions (low FA and high FA). The maximum number of treated wells per medium condition is 880 (16 curves x 11 dilutions x 5 replicates). Complete experimental details and mechanistic implications were reported by Faessel *et al* (18).

3.2. E_{max} model

Due to a plateau of the measure of cell growth such that it does not reach zero at the maximum dose levels used in the experiments, the median effect equation (E4) does not fit the data. Instead, we take the E_{max} model (19) to fit the data at hand.

$$E(d) = E_0 - E_{max} + \frac{E_{max}}{1 + \left(d / ED_{50}\right)^m}, \quad (E5)$$

where E_0 is the base effect, corresponding to the measurement of cell growth when no drug is applied; E_{max} is the maximum effect attributable to the drug; ED_{50} is the dose level producing half of E_{max}; d is the dose level that produces the effect $E(d)$, and m is a slope factor (Hill coefficient) that measures the sensitivity of the effect within a dose range of the drug. Thus, $E_0 - E_{max}$ is the asymptotic effect when a very large dose of the drug is applied. Figure 1 shows a few examples of the E_{max} model where E_0 is assumed to be 1. The parameter m governs how quickly the curve drops. For the three cases in the first row in Figure 1, ED_{50} is fixed at 2 and E_{max} is at 0.8, while the slope varies. When $m=1$ (Figure 1.A), the dose-response

curve drops slowly; when $m=5$ (Figure 1.B and E), a sigmoidal curve is formed, and when $m=20$ (Figure 1.C and F), the drop of the sigmoidal curve becomes very steep. In the three curves in the first row, as the dose increases, the curves drop, and the effect asymptotes to $1 - E_{max} = 0.2$. In the second row, the three plots are set at $E_{max} = 1$, which means that as the dose increases, the treatment will reach the theoretical full effect. For example, if the measure of the treatment effect is cell count, all the cells will be killed at very high doses of the treatment when $E_{max} = 1$. The figures also show that, as ED_{50} increases, the curves shift to the right, indicating that the treatment is less potent. In all cases when m increases, the effect drops more rapidly. We apply the nonlinear weighted least squares method to estimate the parameters in the E_{max} model. Due to the heteroscedascity observed in the data, which means that the variance increases as the observed response increases, we use the reciprocal of the fitted response as the weight function (20). We use S-PLUS, R (21), and SAS (22) to carry out the estimation.

3.3. Interaction index under the E_{max} model

As when using the median effect model, the E_{max} model can be applied to fit the single-drug and combination drug dose-response curves, and then the interaction index can be calculated accordingly. Although equation (E5) allows for different values of E_0 and E_{max} for different curves, when calculating the interaction index, we need to assume all curves have the same E_0 so that the "base measure" of no drug effect is the same in all curves. This can be achieved by dividing all of the effect measures with the mean of the controls. Note that E_{max} can vary in different curves to signify different drug potencies. However, the calculation of the interaction index will be a little more complicated when different drugs or combinations produce different values of E_{max}.

Hereafter, we assume the dose-response curve follows the E_{max} model given in (E6):

$$E(d) = 1 - E_{max} + \frac{E_{max}}{1 + \left(d / ED_{50}\right)^m}. \quad (E6)$$

The experiments we analyzed studied the ability of the combination treatments to inhibit the growth of cancer cells. The measure of the treatment effect was cell growth corresponding to the number of cells observed. Hence, the height of the dose-effect curve decreases when the dose increases. In this case, we have $m > 0$. In addition, as d goes to infinity, the effect plateaus at $1-E_{max}$. Hence, E_{max} must be between 0 and 1.

In a study of two-drug combinations, we need to fit three curves using the E_{max} model: curve 1 for drug 1 alone, curve 2 for drug 2 alone, and curve c for the drug combinations. Denote $E_{max, i}$, $ED_{50, i}$, and m_i as the three parameters for drug i ($i=1,2, c$). Given an effect e ($e>1-E_{max}$), the corresponding dose $d(e)$ can be calculated as
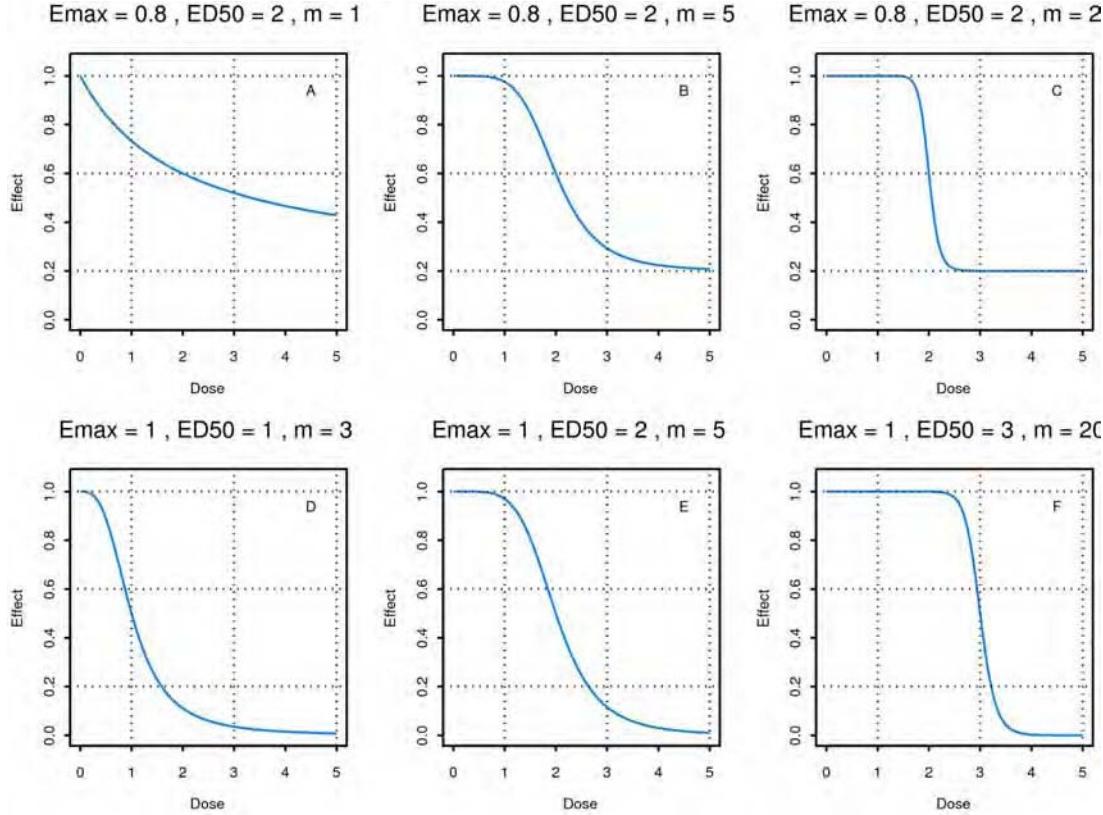
Figure 1. Dose-response curves under the E_{max} model by varying the parameters E_{max}, ED_{50}, and m.

$$d(e) = ED_{50}\left(\frac{1-e}{e-1+E_{max}}\right)^{1/m} \qquad (E7)$$

Note that the dose for the combination treatment is simply the sum of the doses of the single agents. This approach works well for a ray design with constant or varying relative potency between the two drugs (12, 17). Without loss of generality, we can assume that $E_{max,\,1} > E_{max,\,2}$. In addition, we assume that the dose ratio for the two drugs in the combination treatment ($d_c=d_1+d_2$) is fixed with $d_1/\,d_2$ $=p$. Upon fitting the three dose-response curves, the interaction index at a fixed effect e where $e \in (1 - \hat{E}_{max,c}$ be calculated as follows:

$$\hat{II} = \frac{\hat{d}_c(e) \times p/(1+p)}{\hat{D}_{y,1}(e)} + \frac{\hat{d}_c(e)/(1+p)}{\hat{D}_{y,2}(e)} \text{ for } 1-\hat{E}_{max,2} < e < 1, \text{ and}$$

$$\hat{II} = \frac{\hat{d}_c(e) \times p/(1+p)}{\hat{D}_{y,1}(e)} \quad \text{for } 1-\hat{E}_{max,1} < e \leq 1-\hat{E}_{max,2}. \qquad (E8)$$

For $e \leq 1 - \hat{E}_{max,1}$, the interaction index cannot be calculated. However, the combination effect in this range is more than additive because it reaches an effect level that no single agent can achieve alone. If $E_{max,\,1} = E_{max,\,2}$, the interaction index can be calculated using the first formula in ($E8$).

3.4. Confidence interval for the interaction index

We can apply the delta method to calculate the (large sample) variance of the interaction index (23). From our previous work (17), we found that better estimation of

the confidence interval for the interaction index can be achieved by working on the logarithmic transformation of the interaction index.

By applying the delta method, $Var(log(\hat{II})) = \frac{1}{\hat{II}^2}Var(\hat{II})$.

When $1 - \hat{E}_{max,2} < e < 1$, the variance of $\hat{II}$ can be calculated by

$$Var(\hat{II}) = \left(\frac{\hat{d}_c(e) \times p/(1+p)}{\hat{D}_{y,1}(e)}\right)^2 V_1 + \left(\frac{\hat{d}_c(e)/(1+p)}{\hat{D}_{y,2}(e)}\right)^2 V_2 +$$

$$\left(\frac{\hat{d}_c(e) \times p/(1+p)}{\hat{D}_{y,1}(e)} + \frac{\hat{d}_c(e)/(1+p)}{\hat{D}_{y,2}(e)}\right)^2 V_c \qquad (E9)$$

if $1 - \hat{E}_{max,1} < e \leq 1 - \hat{E}_{max,2}$, then

$$Var(\hat{II}) = \left(\frac{\hat{d}_c(e) \times p/(1+p)}{\hat{D}_{y,1}(e)}\right)^2 (V_1 + V_c). \qquad (E10)$$

where

$$V_i = \left(\frac{-1}{\hat{m}_i(e-1+\hat{E}_{max,i})} \quad \frac{1}{\hat{ED}_{50,i}} \quad -\frac{1}{\hat{m}_i^2}log\frac{1-e}{e-1+\hat{E}_{max,i}}\right)Var\left(\hat{E}_{max,i},\hat{ED}_{50,i},\hat{m}_i\right)\begin{pmatrix}\frac{-1}{\hat{m}_i(e-1+\hat{E}_{max,i})}\\[2mm]\frac{1}{\hat{ED}_{50,i}}\\[2mm]-\frac{1}{\hat{m}_i^2}log\frac{1-e}{e-1+\hat{E}_{max,i}}\end{pmatrix}$$

for $i=1, 2, c$.

Upon calculation of the variance for $log(\hat{II})$, the point-wise $(1-\alpha)100\%$ confidence interval for the interaction index (II) for a specified effect can be constructed as

$$\left(\hat{II}exp(-z_{\alpha/2}\sqrt{Var(log(\hat{II}))}), \quad \hat{II}exp(z_{\alpha/2}\sqrt{Var(log(\hat{II}))})\right)(E11)$$

where $z_{\alpha/2}$ is the upper $\alpha/2$ upper percentile of the standard normal distribution. We also construct the simultaneous confidence band for the interaction index over the range of estimated responses. Because the estimation process involves estimating nine parameters from three curves, to construct a Scheffe-type simultaneous confidence band, we simply replace $z_{\alpha/2}$ in equation ($E11$) by $(\text{chi}^2_p(\alpha))^{1/2}$ where $p=9$ (24).

3.5. Data analysis plan

The overall objective of the data analysis is to assess the synergistic effect of the combination of TMQ and AG in both low and high FA media. We apply the exploratory data analysis first, and then estimate the dose-response relationship using the E_{max} model. We evaluate the drug interaction by calculating the interaction index under the Loewe additivity model. We perform an exploratory data analysis in order to understand the data structure and patterns and to determine whether preprocessing of the data in terms of outlier rejection and standardization would be required prior to data modeling. We analyze the low FA and high FA experiments separately then compare the results. For each experiment, we apply the E_{max} model to fit the two marginal and twelve combination dose-response curves. We compute the interaction index and its 95% confidence intervals for each of the twelve combinations, and assess the overall pattern of drug interaction by examining the interaction index from the 12 fixed-ratio combinations together. We apply a one-dimensional distribution plot via the BLiP plot (25) to display the data. We use a two-dimensional scatter plot, a contour plot, and an image plot as well as a three-dimensional perspective plot to show the dose-response relationship. We also apply a trellis plot (26) to assemble the individual plots together into consecutive panels conditioning on different values of fixed dose ratios.

4. EXPLORATORY DATA ANALYSIS

As in all data analyses, we begin with an exploratory data analysis. For the low and high FA experiments, there are 871 and 879 readings, respectively. Only 9 and 1 observations, respectively, are missing out of a maximum of 880 readings in each experiment. The data include designated curve numbers ranging from 1 to 16 and data point numbers ranging from 1 to 176. Each curve number indicates a specific dose combination. We re-label the curves as A-P where A and B correspond to the control (no drug) curves; C and D correspond to the curves of TMQ and AG administered alone, and curves E through P correspond to the combination curves with fixed dose ratios in ascending order. Each point number indicates the readings at each specific dilution of each curve. Because five duplicated experiments were performed, there are up to five readings for each specific point number. There is, however, no designation of the plate number in the data received. Figure 2 shows the variable percentile plot of the distribution of the effect from the low FA and high FA experiments using the BLiP plot, with each segment corresponding to a five percent increment (25). The plot gives an overall

assessment of the distribution of the outcome variable of cell growth without conditioning on experimental settings. The middle 20% of the data (40th to 60th percentiles) are shaded in a light orange color. This figure indicates that the data have a bimodal distribution with most data clustered around either a low value of 0.2 or a high value of 1.2. For the low FA experiment, the distribution of the effect ranges from 0.072 to 1.506 with the lower, middle, and upper quartiles being 0.149, 0.449, and 1.150, respectively. Similarly, for the high FA experiment, the effect ranges between 0.070 and 1.545. The three respective quartiles are 0.213, 0.990, and 1.1495. The median of the data from the low FA experiment is smaller than the median of the data from the high FA experiment. The bimodal distributions could result from steep dose-response curves. As a consequence, the slope may not be estimated well in certain cases.

To help understand the pattern of the fixed ratio dose assignment in a ray design and the relationship between the fixed ratio doses and curve numbers, we plot the logarithm transformed dose of TMQ and AG in Figure 3 for both the low FA and high FA experiments. As can be seen, curves A and B are the controls with no drugs. Curves C and D correspond to the single drug study of TMQ and AG, respectively. Curves E through P are the various fixed ratio combination doses of TMQ and AG. Note that curves J and K have the same dose ratios. Within each curve, the 11 dilutions are marked by 11 circles. For the combination studies, the curves for different dose ratios are parallel to each other on the log dose scale. If the same plot is shown in the original scale, these lines will form "rays," radiating out from the origin like sun rays. Hence, the term "ray design" is an appropriate name for this type of experiment. The corresponding dose ranges used for each drug alone are 5.47×10^{-6} to 0.56 µM for TMQ in both the low FA and high FA experiments, and 2.71×10^{-5} to 2.78 µM for AG2034 in the low FA experiment and 2.71×10^{-4} to 27.78 µM in the high FA experiment.

Figures 4 and 5 show the raw data of the effect versus dose level by curve for the low FA and high FA experiments, respectively. Instead of using the actual dose, we plot the data using a sequentially assigned dose level to indicate each dilution within each curve such that the data can be shown clearly. In addition, the data points at each dilution for each curve are coded from 1 to 5 according to the order of the appearance in the data set. We assume that these numbers correspond to the replicate number for each design point (the well position in the stack of 5 plates). Because the plate number was not listed in the data, we are not certain that this is the case. From the plot, we can see that there are outliers in several dilution series. Of note, in Figure 4, the effects from plate (replicate) #1 in curves B, E, F, and K tend to be lower than all other replicates. There are also some unusually large values, for example, in replicate 2 in curve A, dose level (dilution series) 6; replicate 3 in curve L, dose level 4; and replicate 2 in curve M, dose level 1. Similar observations can be made for the
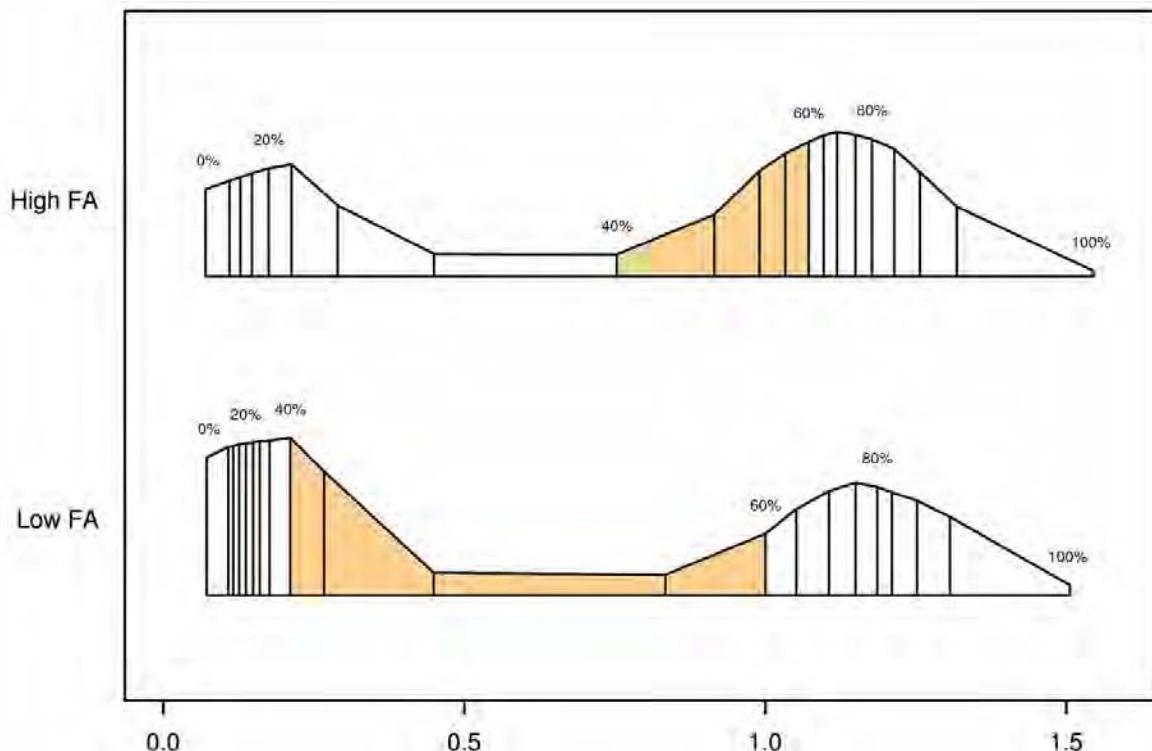
Figure 2. Variable width percentile plot for the observed effect in experiments with low and high folic acid media. Each vertical bar indicates a five percent increment. The middle 20% of the data are shaded in a light orange color.

high FA experiment: plate #1 seems to have some low values in curves B, C, H, I, and J, and plate #4 seems to have some low values in curves E, K, N, O, and P. These findings indicate that certain procedures need to be performed to remove the obvious outliers in order to improve the data quality before the data analysis.

Figure 6 shows the perspective plot, contour plot, and image plot for the low FA experiment. From the perspective plots in Figure 6.A (back view), B (front view), and C (side view), we can see that the effect starts at a high plane plateau at an effect level of about 1.2 when the doses of TMQ are AG are small. As the dose of each drug increases, the effect remains approximately constant for a while and then a sudden drop occurs. This steep downward slope can be found by taking the trajectory of any combination of the TMQ and AG doses; it is also evident in the dose-response curves shown in Figures 4 and 5. The steep drop of the effect can also be found in the contour plot and the image plot. Similar patterns in the dose-response relationship are shown in Figure 7 for the high FA experiment. The steep drop of the effect occurs at smaller doses in the low FA experiment and at larger doses in the high FA experiment.

5. DATA PREPROCESSING: OUTLIER REJECTION AND DATA STANDARDIZATION

5.1. Outlier rejection

To address the concern that outliers may adversely affect the analysis outcome, we devise the following simple plan. For each of the 176 point numbers (16 curves x 11 dilutions), the five effect readings should be close to each other because they are from replicated experiments. However, because the plate number is not in the data set, we cannot assess the plate effect. Neither can we reject a certain replicate plate entirely should there be a plate with outlying data, nor apply a mixed effect model treating the plate effect as a random effect. For the four or five effect readings in each point number (only 9 point numbers in the low FA and 1 in the high FA experiments have 4 readings), we compute the median and the interquartile range. An effect reading is considered an outlier if the value is beyond the median ± 1.4529 times the interquartile range. If the data are normally distributed (i.e., follow a Gaussian distribution), the range expands to cover the middle 95% of the data. Hence, only about 5% of the data points (2.5% at each extreme) are considered outliers. The number 1.4529 is obtained by qnorm(.975)/(qnorm(.75) - qnorm(.25)) where qnorm(x) is a quantile function which returns the xth percentiles from a normal distribution. Upon applying the above rule, 129 out of 871 (14.8%) effect readings in the low FA experiment and 126 out of 879 (14.3%) in the high FA experiment are considered outliers and are removed before proceeding to further analysis. The numbers of outliers in replicates 1 to 5 are 60, 28, 19, 14, and 8 for the low FA experiment and 35, 18, 21, 34, and 18 for the high FA experiment, indicating a non-random pattern of outliers that could be attributed to experimental conditions. Note that the outlier rejection algorithm is only applied "locally." In other words, it only applies to the replicated readings up to five replicates in each of the 176 experimental conditions.
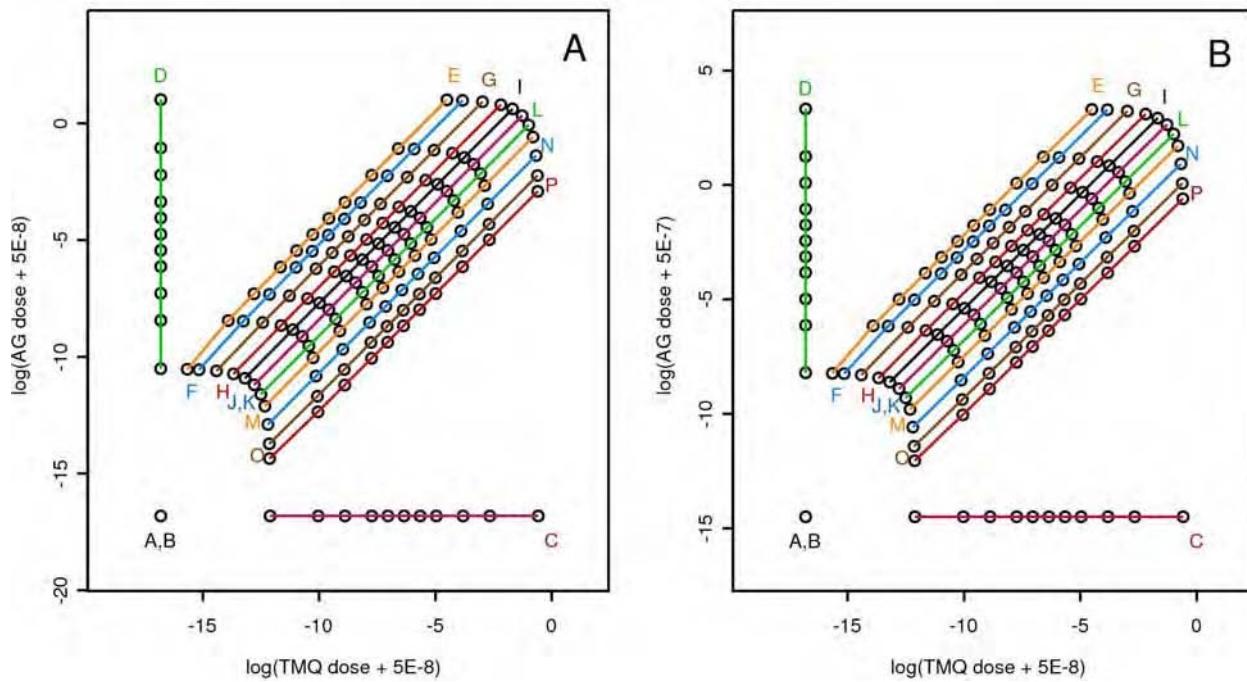
Figure 3. Experimental design showing the logarithmically transformed AG2034 (AG) dose versus the logarithmically transformed trimetrexate (TMQ) dose in the fixed ratio experiments. 16 curves are shown. Curves A and B are controls; no drugs applied. Curves C and D are single-drug studies for TMQ and AG, respectively. Curves E through P are the combination drug studies. Each curve has 11 dilutions shown in circles. Panel A: low folic acid medium. Panel B: high folic acid medium.

5.2. Data standardization

After outliers are removed from the data, we compute the mean of the control curves. The means for curves 8 and 16 are 1.1668 and 1.1534 for the low FA experiment and 1.1483 and 1.1477 for the high FA experiment, respectively. To apply the E_{max} model in equation ($E6$) with $E_0 = 1$, we standardize the data by dividing the effect readings of respective curves 1-7 by the mean of curve 8 and the effect readings of respective curves 9-15 by the mean of curve 16.

6. RESULTS

6.1. Results for the low folic acid experiment

The E_{max} model in equation ($E6$) is applied to fit all of the dose-response curves. For the low FA experiment, the parameter estimates, their corresponding standard errors, and the residual sum of squares are given in Table 1. The dose-response relationships showing the data and the fitted curves are displayed in Figure 8. Note that although model fitting is performed on the original dose scale, the dose is plotted on the logarithmically transformed scale to better show the dose-response relationship. The fitted marginal dose-response curves for TMQ (curve C) and AG (curve D) are shown in a blue dashed line and a red dotted line, respectively. From Table 1, we see that $\hat{ED}_{50}$ is 0.00133 for TMQ and 0.00621 for AG, indicating that TMQ is about 4.7 times more potent than AG at the ED_{50} level. For

curves E through P, the fitted dose-response curve for the combination treatment is shown as a solid black line superimposed on the marginal dose-response curves. The proposed E_{max} model fits all curves well except for curves G, H and K. For curve G, although the model estimates converge in an initial attempt, the parameter m is estimated with a standard error of 30.3. The large standard error essentially indicates that the estimate $\hat{m}$ is not reliable. For curve K, the model does not converge on the original dose scale but converges on the logarithmically transformed dose scale. However, the standard error of the estimate $\hat{m}$ is still very large, which leads us to believe that the model is not very stable. For curve H, as can be seen in Figure 8, there are no observed effects between 0.3 and 1 from the second to the fifth dilutions. The parameter m cannot be estimated and the model fails to converge on both the original scale and the logarithmic scale. To address these problems, we conclude that the data do not provide us sufficient information to yield a reasonable estimate of the parameter m. Therefore, we take a remedial approach by fixing m, and then proceed to estimate the other two parameters. Upon checking the data, we set the parameter m as 5, 4.5, and 5 for curves G, H, and K, respectively. The choice of m is somewhat arbitrary with a goal of yielding a good fit to the data and producing a small residual sum of squares. The resulting "reduced" models fit the data reasonably well but with a consequence that there is no standard error estimate for $\hat{m}$, which affects the variance estimation of the interaction index (to be shown later).

Table 1. Summary of parameter estimates (standard error) for the low FA experiment

Curve[1]	Dose ratio (TMQ/AG)[2]	E_{max}[3]	ED_{50}[4]	m[5]	Residual sum of squares
C (1)		0.877 (0.007)	0.00133 (0.00006)	2.345 (0.190)	0.0779
D (2)		0.872 (0.007)	0.00621 (0.00024)	3.045 (0.269)	0.0749
E (15)	0.004	0.869 (0.008)	0.00359 (0.00017)	3.250 (0.437)	0.0969
F (13)	0.008	0.863 (0.008)	0.00294 (0.00014)	2.621 (0.276)	0.0897
G (11*)	0.02	0.865 (0.006)	0.00151 (0.00005)	5.0	0.0817
H (7*)	0.05	0.889 (0.007)	0.00274 (0.00011)	4.5	0.1025
I (5)	0.1	0.885 (0.005)	0.00253 (0.00009)	3.449 (0.306)	0.0689
J (3)	0.2	0.882 (0.005)	0.00244 (0.00007)	4.019 (0.402)	0.0655
K (9*)	0.2	0.872 (0.007)	0.00233 (0.00007)	5.0	0.0843
L (4)	0.4	0.889 (0.006)	0.00278 (0.00011)	5.473 (0.583)	0.0855
M (6)	0.8	0.890 (0.005)	0.00200 (0.00007)	3.208 (0.263)	0.0738
N (10)	2	0.887 (0.008)	0.00169 (0.00009)	2.544 (0.258)	0.0984
O (12)	5	0.878 (0.009)	0.00145 (0.00007)	2.206 (0.206)	0.0837
P (14)	10	0.874 (0.006)	0.00134 (0.00006)	1.971 (0.128)	0.0599

Footnotes and abbreviations: [1]curves without data: curves A(8) and B(16) represent controls; no drugs applied, [2]trimetrexate/experimental drug AG2034, [3]maximum effect attributable to the drug, [4]dose level producing half of E_{max}, [5]slope factor, which measures the sensitivity of the effect within a dose range of the drug, is fixed at a certain value

Based on limited sensitivity analysis, the estimation of the interaction index remains reasonably robust.

In all dose-response curves, the standardized effect level starts to drop between dose levels (dilutions) 3 to 6. Once the effect starts to drop, it drops quickly and plateaus at the $1 - \hat{E}_{max}$ level. There are ample data points at the effect levels around 1 (dose levels 1-4) and $1 - \hat{E}_{max}$ (dose levels 8-11). However, due to the sharp drop in the dose-response curves, fewer data points can be found in the middle of the effect range. When the data points become too few or do not spread out to cover enough range, it becomes harder for the model to converge, as seen in curves G, H, and K. The overall results for the curve fitting of the low FA experiment are that the values of $\hat{E}_{max}$ range from 0.863 to 0.890; the values of $\hat{ED}_{50}$ range from 0.00133 to 0.00621; and the values of $\hat{m}$ range from 1.971 to 5.473. The residual sum of squares ranges from 0.0599 to 0.1025 without large values, suggesting that the model fits the data reasonably well.

Based on the fitted dose-response curve, the interaction index (II) can be calculated over the entire effect range and at specific dose combinations. Table 2 gives a detailed result of the estimated interaction index and its 95% point-wise confidence interval at each dose combination for each combination curve. The II is calculated at the predicted effect level from the combination curve and not at the observed effect level. The results are shown in a trellis plot in Figure 9 where the red lines represent the 95% point-wise confidence intervals at each specific effect level and the black dashed lines indicate the 95% simultaneous confidence bands of the II for the entire range. From the figure we find that the interaction index can be estimated with very good precision in all curves except at the two extremes when the effect is close to 1 or $1 - \hat{E}_{max}$. The trend and the pattern of the interaction index are clearly shown in these figures. For curves E through K, i.e., with a TMQ:AG dose ratio ranging from 0.004 to 0.2, synergy is observed in the effect range between 0.2 to 0.9. For curves L and M, which have

TMQ:AG ratios of 0.4 and 0.8, we see that synergy is observed at the low effect level from 0.2 to about 0.5. Beyond 0.5 the combinations are generally additive. For curves N, O, and P, with TMQ:AG ratios of 2, 5, and 10, the synergistic effect is lost and we see additivity in all dose ranges.

6.2. Results for the high folic acid experiment

Table 3 lists the parameter estimate, corresponding standard error, and sum of squares for all the curves in the high FA experiment. Unlike in the experiment using low FA media, the model fitting for all curves in the high FA experiment converge when using the E_{max} model. The estimated $\hat{E}_{max}$ ranges from 0.831 to 0.893; $\hat{ED}_{50}$ ranges from 0.0137 to 0.1943 except for curve D (AG alone with $\hat{ED}_{50} = 0.5224$); and $\hat{m}$ ranges between 1.468 and 3.625. The residuals sum of squares ranges from 0.0615 to 0.1134. Compared to the low FA experiment, the values of $\hat{ED}_{50}$ are greater in the high FA experiment, indicating that the drugs are less potent when applied to a high FA medium. Note that the doses of TMQ are the same between the two experiments but the doses of AG are 10 times higher in the high FA experiment. In addition, $\hat{ED}_{50} = 0.0137$ and 0.00133 for TMQ alone in the high and low FA experiments, respectively, which indicates that the drug is 10 times less potent in the high FA medium compared to the low FA medium. The potency of AG is even more dramatically reduced. In Figure 10 we see that the E_{max} model provides an excellent fit to all the curves. Table 4 gives a detailed account of the interaction index in all dilutions for all of the combination curves. The results are summarized in a trellis plot in Figure 11. Again, the red lines represent the 95% point-wise confidence intervals at each specific effect level and the black dashed lines correspond to the simultaneous confidence bands of the II for the whole range. Using the high FA medium, synergy can be achieved for most of the drug combinations in all the effect ranges, with the exception of the very low or very high effect ranges. The confidence intervals are still very tight although they are a little wider compared to their counterparts from the low FA experiment. As the TMQ:AG

Table 2. Estimated interaction index and its 95% confidence interval at each dose combination for the low FA experiment

Curve[1]	TMQ[2] dose	AG2034[3] dose	Dose ratio TMQ/AG	Dilution	Predicted effect	II[4]	95% CI[5] for II Lower limit	Upper limit
E (15)	1.07E-07	2.66E-05	0.004	1	1	0.87	0.18	4.29
	8.58E-07	0.000213		2	0.9999	0.73	0.28	1.89
	2.71E-06	0.000673		3	0.9962	0.67	0.37	1.20
	8.58E-06	0.002129		4	0.864	0.61	0.48	0.78
	1.72E-05	0.004259		5	0.4454	0.58	0.52	0.65
	3.43E-05	0.008517		6	0.1802	0.56	0.45	0.71
	6.86E-05	0.017000		7	0.1368	0.61	0.24	1.55
	0.000137	0.034100		8	0.1319	0.91	0	4.35E+03
	0.000434	0.107700		9	0.1314	2.71	0	1.30E+157
	0.001373	0.340700		10	0.1314	8.58	0	NA
	0.011000	2.725500		11	0.1314	68.6	0	NA
F (13)	2.10E-07	2.61E-05	0.008	1	1	0.28	0.08	0.91
	1.68E-06	0.000209		2	0.9991	0.35	0.17	0.71
	5.32E-06	0.000660		3	0.9828	0.4	0.26	0.62
	1.68E-05	0.002088		4	0.746	0.47	0.4	0.55
	3.37E-05	0.004177		5	0.3788	0.52	0.46	0.58
	6.73E-05	0.008353		6	0.188	0.59	0.47	0.74
	0.000135	0.016700		7	0.1454	0.76	0.4	1.47
	0.000269	0.033400		8	0.138	1.26	0.02	67.11
	0.000851	0.105700		9	0.1366	3.79	0	1.72E+37
	0.002692	0.334100		10	0.1366	11.95	0	NA
	0.021500	2.673100		11	0.1366	95.63	0	NA
G (11)	4.97E-07	2.47E-05	0.02	1	1.0000	5.60	1.90	16.45
	3.98E-06	0.000197		2	1.0000	1.09	0.60	1.96
	1.26E-05	0.000624		3	0.9885	0.47	0.36	0.63
	3.98E-05	0.001974		4	0.2987	0.22	0.19	0.25
	7.95E-05	0.003949		5	0.1410	0.17	0.09	0.32
	0.000159	0.007898		6	0.1350	0.27	0.00	3.09E+04
	0.000318	0.015800		7	0.1348	0.54	0.00	1.62E+161
	0.000636	0.031600		8	0.1348	1.09	0.00	NA
	0.002012	0.099900		9	0.1348	3.44	0.00	NA
	0.006364	0.315900		10	0.1348	10.87	0.00	NA
	0.050900	2.527300		11	0.1348	86.95	0.00	NA
H (7)	1.09E-06	2.17E-05	0.05	1	1.0000	12.11	3.53	41.51
	8.75E-06	0.000174		2	1.0000	2.71	1.42	5.15
	2.77E-05	0.000549		3	0.9992	1.29	0.88	1.90
	8.75E-05	0.001738		4	0.8773	0.65	0.57	0.74
	0.000175	0.003475		5	0.3035	0.43	0.38	0.48
	>=0.000350	>=0.006950		6 - 11	<= 0.1219	NA	NA	NA
I (5)	1.82E-06	1.81E-05	0.1	1	1.0000	2.39	0.70	8.17
	1.46E-05	0.000145		2	0.9999	1.17	0.59	2.30
	4.61E-05	0.000458		3	0.9966	0.83	0.55	1.24
	0.000146	0.001448		4	0.8509	0.61	0.52	0.71
	0.000292	0.002896		5	0.3906	0.51	0.47	0.55
	0.000583	0.005792		6	0.1506	0.38	0.31	0.47
	>=0.001167	>=0.011600		7 - 11	<= 0.1188	NA	NA	NA
J (3)	2.73E-06	1.36E-05	0.2	1	1.0000	12.56	2.59	60.87

	2.19E-05	0.000109		2	1.0000	3.31	1.38	7.95
	6.92E-05	0.000343		3	0.9993	1.67	0.99	2.81
	0.000219	0.001086		4	0.9344	0.88	0.71	1.08
	0.000438	0.002172		5	0.5008	0.61	0.57	0.65
	0.000875	0.004344		6	0.1577	0.40	0.33	0.48
	>=0.001750	>=0.008688		7 - 11	<= 0.1204	NA	NA	NA
K (9)	2.73E-06	1.36E-05	0.2	1	1.0000	88.77	14.80	532.45
	2.19E-05	0.000109		2	1.0000	9.70	3.84	24.48
	6.92E-05	0.000343		3	0.9998	3.04	1.83	5.06
	0.000219	0.001086		4	0.9550	1.02	0.86	1.21
	0.000438	0.002172		5	0.4457	0.55	0.51	0.60
	0.000875	0.004344		6	0.1429	0.32	0.21	0.47
	>=0.001750	>=0.008688		7 - 11	<= 0.1280	NA	NA	NA
L (4)	3.65E-06	9.05E-06	0.4	1	1.0000	812.88	82.77	7.98E+03
	2.92E-05	7.24E-05		2	1.0000	53.71	13.70	210.53
	9.22E-05	0.000229		3	1.0000	12.38	5.18	29.56
	0.000292	0.000724		4	0.9964	2.99	1.98	4.51
	0.000583	0.001448		5	0.8651	1.31	1.10	1.56
	0.001167	0.002896		6	0.2103	0.56	0.50	0.63
	>=0.002333	>=0.005792		7 - 11	<= 0.1134	NA	NA	NA
M (6)	4.38E-06	5.43E-06	0.8	1	1.0000	4.95	1.18	20.69
	3.50E-05	4.34E-05		2	1.0000	2.40	1.01	5.71
	0.000111	0.000137		3	0.9989	1.63	0.93	2.87
	0.000350	0.000434		4	0.9580	1.12	0.85	1.48
	0.000700	0.000869		5	0.7206	0.90	0.80	1.03
	0.001400	0.001738		6	0.2804	0.72	0.64	0.80
	0.002800	0.003475		7	0.1325	0.41	0.25	0.67
	>=0.005600	>=0.006950		8 - 11	<= 0.1128	NA	NA	NA
N (10)	4.97E-06	2.47E-06	2	1	1.0000	1.37	0.32	5.84
	3.98E-05	1.97E-05		2	0.9998	1.17	0.47	2.89
	0.000126	0.000062		3	0.9967	1.08	0.59	1.98
	0.000398	0.000197		4	0.9417	0.99	0.72	1.37
	0.000795	0.000395		5	0.7418	0.95	0.80	1.13
	0.001591	0.000790		6	0.3742	0.90	0.79	1.02
	0.003182	0.001580		7	0.1721	0.81	0.64	1.03
	>=0.006364	>=0.003159		8 - 11	<= 0.1236	NA	NA	NA
O (12)	5.26E-06	1.04E-06	5	1	1.0000	0.67	0.18	2.52
	4.21E-05	8.35E-06		2	0.9995	0.76	0.33	1.76
	0.000133	2.64E-05		3	0.9934	0.82	0.46	1.45
	0.000421	8.35E-05		4	0.9227	0.88	0.65	1.21
	0.000841	0.000167		5	0.7294	0.92	0.78	1.10
	0.001683	0.000334		6	0.4094	0.97	0.87	1.08
	0.003365	0.000668		7	0.2060	1.01	0.83	1.24
	0.006731	0.001337		8	0.1420	1.05	0.67	1.65
	>=0.021300	>=0.004227		9 - 11	<= 0.1239	NA	NA	NA
P (14)	5.36E-06	5.32E-07	10	1	1.0000	0.39	0.14	1.09
	4.29E-05	4.26E-06		2	0.9988	0.54	0.28	1.05
	0.000136	1.35E-05		3	0.9887	0.65	0.42	1.02
	0.000429	4.26E-05		4	0.9015	0.79	0.62	1.01

0.000858	8.52E-05	5	0.7095	0.88	0.76	1.02
0.001716	0.000170	6	0.4221	0.99	0.89	1.10
0.003431	0.000341	7	0.2272	1.12	0.95	1.32
0.006863	0.000681	8	0.1544	1.30	0.98	1.74
>=0.021700	>=0.002155	9 - 11	<= 0.1292	NA	NA	NA

Footnotes and abbreviations: [1] curves without data: curves A (8) and B (16) represent controls; no drugs applied; curves C (1, TMQ) and D (2, AG2034) represent single-drug applications, [2] trimetrexate, [3] experimental drug, [4] interaction index, [5] confidence interval

Table 3. Summary of parameter estimates (standard error) for the high FA experiment

Curve[1]	Dose ratio (TMQ/AG)[2]	E_{max}[3]	ED_{50}[4]	m[5]	Residual sum of squares
C (1)		0.883 (0.012)	0.0137 (0.0012)	3.625 (0.650)	0.1074
D (2)		0.831 (0.015)	0.5224 (0.0439)	1.468 (0.137)	0.0770
E (15)	0.0004	0.867 (0.014)	0.1943 (0.0122)	2.558 (0.405)	0.1134
F (13)	0.0008	0.863 (0.010)	0.1447 (0.0068)	2.643 (0.258)	0.0852
G (11)	0.002	0.859 (0.006)	0.0912 (0.0045)	2.996 (0.355)	0.0999
H (7)	0.005	0.881 (0.006)	0.0699 (0.0027)	2.887 (0.253)	0.0746
I (5)	0.01	0.881 (0.009)	0.0484 (0.0026)	2.528 (0.251)	0.0977
J (3)	0.02	0.884 (0.006)	0.0331 (0.0011)	2.114 (0.136)	0.0615
K (9)	0.02	0.885 (0.008)	0.0369 (0.0019)	2.160 (0.195)	0.0861
L (4)	0.04	0.886 (0.008)	0.0288 (0.0014)	2.504 (0.255)	0.0959
M (6)	0.08	0.885 (0.009)	0.0197 (0.0010)	2.242 (0.214)	0.0881
N (10)	0.2	0.862 (0.010)	0.0154 (0.0007)	3.309 (0.415)	0.0909
O (12)	0.5	0.878 (0.009)	0.0139 (0.0006)	3.491 (0.405)	0.0933
P (14)	1	0.893 (0.008)	0.0183 (0.0009)	2.735 (0.213)	0.0669

Footnotes and abbreviations: [1] curves without data: curves A(8) and B(16) represent controls; no drugs applied, [2] trimetrexate/experimental drug AG2034, [3] maximum effect attributable to the drug, [4] dose level producing half of E_{max}, [5] slope factor, which measures the sensitivity of the effect within a dose range of the drug

Table 4. Estimated interaction index and its 95% confidence interval at each dose combination for the high FA experiment

Curve[1]	TMQ[2] dose	AG2034[3] dose	Dose ratio TMQ/AG	Dilution	Predicted effect	II[4]	95% CI[5] for II Lower limit	Upper limit
E (15)	1.07E-07	0.000266	0.0004	1	1.0000	48.28	2.53	922.71
	8.58E-07	0.002128		2	1.0000	10.31	1.36	78.32
	2.71E-06	0.006729		3	0.9998	4.39	0.96	20.02
	8.58E-06	0.021278		4	0.9970	1.87	0.68	5.14
	1.72E-05	0.042555		5	0.9825	1.12	0.55	2.28
	3.43E-05	0.085110		6	0.9063	0.67	0.44	1.02
	6.86E-05	0.170221		7	0.6388	0.40	0.32	0.48
	0.000137	0.340441		8	0.2994	0.21	0.16	0.28
	>=0.000434	>=1.076570		9 - 11	<= 0.1433	NA	NA	NA
F (13)	2.10E-07	0.000261	0.0008	1	1.0000	42.38	3.74	479.71
	1.68E-06	0.002087		2	1.0000	8.02	1.56	41.25
	5.32E-06	0.006599		3	0.9998	3.20	0.96	10.63
	1.68E-05	0.020868		4	0.9949	1.27	0.59	2.75
	3.37E-05	0.041737		5	0.9688	0.73	0.44	1.23
	6.73E-05	0.083474		6	0.8363	0.42	0.32	0.56
	0.000135	0.166947		7	0.4876	0.24	0.20	0.28
	0.000269	0.333894		8	0.2224	0.11	0.08	0.16
	>=0.000851	>=1.055866		9 - 11	<= 0.1418	NA	NA	NA
G (11)	4.97E-07	0.000247	0.002	1	1.0000	80.07	5.84	1097.54
	3.98E-06	0.001973		2	1.0000	9.20	1.67	50.66
	1.26E-05	0.006239		3	0.9997	2.78	0.84	9.26
	3.98E-05	0.019730		4	0.9913	0.85	0.42	1.71
	7.95E-05	0.039460		5	0.9351	0.42	0.28	0.63
	0.000159	0.078920		6	0.6609	0.21	0.17	0.25

Emax model and interaction index for drug interaction

	0.000318	0.157841		7	0.2796	0.10	0.08	0.12

	0.000318	0.157841		7	0.2796	0.10	0.08	0.12
	>=0.000636	>=0.315682		8 - 11	<= 0.1614	NA	NA	NA
H (7)	1.09E-06	0.000217	0.005	1	1.0000	33.80	3.34	342.20
	8.75E-06	0.001736		2	1.0000	4.54	1.03	20.08
	2.77E-05	0.005491		3	0.9994	1.50	0.54	4.20
	8.75E-05	0.017363		4	0.9843	0.51	0.29	0.90
	0.000175	0.034725		5	0.8955	0.27	0.20	0.37
	0.000350	0.069450		6	0.5603	0.15	0.13	0.17
	0.000700	0.138900		7	0.2239	0.07	0.06	0.09
	>=0.001400	>=0.277800		8 - 11	<= 0.1344	NA	NA	NA
I (5)	1.82E-06	0.000181	0.01	1	1.0000	4.94	0.61	39.76
	1.46E-05	0.001447		2	0.9999	1.11	0.30	4.13
	4.61E-05	0.004575		3	0.9977	0.50	0.21	1.20
	0.000146	0.014469		4	0.9592	0.23	0.15	0.37
	0.000292	0.028938		5	0.8071	0.16	0.12	0.20
	0.000583	0.057875		6	0.4556	0.11	0.09	0.13
	0.001167	0.115750		7	0.2041	0.07	0.06	0.09
	>=0.002333	>=0.231500		8 - 11	<= 0.1347	NA	NA	NA
J (3)	2.73E-06	0.000136	0.02	1	1.0000	0.67	0.13	3.34
	2.19E-05	0.001085		2	0.9993	0.28	0.10	0.74
	6.92E-05	0.003432		3	0.9924	0.18	0.09	0.34
	0.000219	0.010852		4	0.9206	0.13	0.09	0.17
	0.000438	0.021703		5	0.7354	0.11	0.09	0.13
	0.000875	0.043406		6	0.4261	0.10	0.09	0.12
	0.001750	0.086813		7	0.2139	0.10	0.08	0.11
	>=0.003500	>=0.173625		8 - 11	<= 0.1405	NA	NA	NA
K (9)	2.73E-06	0.000136	0.02	1	1.0000	0.93	0.15	5.72
	2.19E-05	0.001085		2	0.9995	0.36	0.12	1.12
	6.92E-05	0.003432		3	0.9946	0.22	0.11	0.47
	0.000219	0.010852		4	0.9390	0.15	0.10	0.22
	0.000438	0.021703		5	0.7800	0.13	0.10	0.16
	0.000875	0.043406		6	0.4722	0.11	0.10	0.13
	0.001750	0.086813		7	0.2316	0.11	0.09	0.13
	>=0.003500	>=0.173625		8 - 11	<= 0.1443	NA	NA	NA
L (4)	3.65E-06	0.000090	0.04	1	1.0000	2.89	0.34	24.37
	2.92E-05	0.000723		2	0.9999	0.69	0.18	2.62
	9.22E-05	0.002288		3	0.9983	0.33	0.14	0.80
	0.000292	0.007234		4	0.9702	0.18	0.12	0.29
	0.000583	0.014469		5	0.8538	0.15	0.11	0.19
	0.001167	0.028938		6	0.5312	0.13	0.11	0.15
	0.002333	0.057875		7	0.2326	0.12	0.10	0.15
	>=0.004667	>=0.115750		8 - 11	<= 0.1355	NA	NA	NA
M (6)	4.38E-06	5.43E-05	0.08	1	1.0000	0.73	0.10	5.17
	3.50E-05	0.000434		2	0.9998	0.27	0.08	0.89
	0.000111	0.001373		3	0.9973	0.17	0.08	0.37
	0.000350	0.004341		4	0.9660	0.13	0.09	0.21
	0.000700	0.008681		5	0.8594	0.13	0.10	0.17
	0.001400	0.017363		6	0.5823	0.14	0.12	0.16
	0.002800	0.034725		7	0.2842	0.16	0.13	0.18
	>=0.005600	>=0.069450		8 - 11	<= 0.1571	NA	NA	NA

N (10)	4.97E-06	2.47E-05	0.2	1	1.0000	61.54	3.00	1262.99
	3.98E-05	0.000197		2	1.0000	4.66	0.64	34.02
	0.000126	0.000624		3	1.0000	1.21	0.31	4.73
	0.000398	0.001973		4	0.9983	0.41	0.20	0.87
	0.000795	0.003946		5	0.9830	0.28	0.17	0.46
	0.001591	0.007892		6	0.8570	0.23	0.17	0.31
	0.003182	0.015784		7	0.4280	0.21	0.18	0.25
	0.006364	0.031568		8	0.1800	0.23	0.18	0.30
	>=0.020124	>=0.099827		9 - 11	<= 0.1390	NA	NA	NA
O (12)	5.26E-06	1.04E-05	0.5	1	1.0000	194.98	6.90	5509.68
	4.21E-05	8.35E-05		2	1.0000	11.38	1.16	111.41
	0.000133	0.000264		3	1.0000	2.57	0.51	12.90
	0.000421	0.000835		4	0.9998	0.78	0.30	1.97
	0.000841	0.001669		5	0.9978	0.50	0.25	0.98
	0.001683	0.003339		6	0.9754	0.40	0.24	0.64
	0.003365	0.006678		7	0.7849	0.36	0.27	0.48
	0.006731	0.013356		8	0.3109	0.35	0.30	0.41
	>=0.021285	>=0.042235		9 - 11	<= 0.1263	NA	NA	NA
P (14)	5.36E-06	5.32E-06	1	1	1.0000	10.38	0.65	165.28
	4.29E-05	4.26E-05		2	1.0000	1.89	0.29	12.19
	0.000136	0.000135		3	1.0000	0.87	0.24	3.15
	0.000429	0.000426		4	0.9998	0.55	0.23	1.29
	0.000858	0.000851		5	0.9986	0.50	0.25	1.03
	0.001716	0.001702		6	0.9910	0.51	0.29	0.92
	0.003431	0.003404		7	0.9436	0.56	0.37	0.86
	0.006863	0.006809		8	0.7232	0.64	0.50	0.83
	0.021702	0.021531		9	0.1851	0.80	0.65	0.99
	>=0.068627	>=0.068088		10, 11	<= 0.1109	NA	NA	NA

Footnotes and abbreviations: [1]curves without data: curves A(8) and B(16) represent controls; no drugs applied; curves C(1, TMQ) and D (2, AG2034) represent single-drug applications, [2]trimetrexate, [3]experimental drug, [4]interaction index, [5]confidence interval

ratio increases from 0.0004 to 0.5, synergy is observed across all dilution series. In addition, higher synergy is observed at the lower effect levels, particularly when the TMQ:AG is at 0.01 or lower (curves E, F, G, H, and I). In the middle effect levels (effects between 0.2 and 0.8), the II ranges from about 0.1 in curves J and K, to 0.12 in curve L, 0.15 in curve M, 0.25 in curve N, and 0.35 in curve O. The higher the TMQ:AG ratio, the less synergy it achieves. In curve P, for example, when the TMQ:AG ratio reaches 1, synergy is lost.

7. SUMMARY

In both the low FA and high FA experiments, TMQ is more potent than AG. At low TMQ:AG ratios, i.e., when a small amount of the more potent drug (TMQ) is added to a larger amount of the less potent drug (AG), synergy is achieved. However, when the TMQ:AG ratio reaches 0.4 or larger for the low FA medium, or when the TMQ:AG ratio reaches 1 or larger for the high FA medium, synergy decreases, or the interaction becomes additive. In general, a synergistic effect in a drug combination dilution series is stronger at higher doses that produce stronger effects (effects closer to $1-E_{max}$) than at lower dose levels

that produce weaker effects (effects closer to 1). The two drugs in this study are more potent in the low FA medium compared to the high FA medium. The drug synergy, however, is stronger in the high FA medium.

8. DISCUSSION AND PERSPECTIVE

The data supplied by Dr. Greco provide an excellent opportunity to apply and compare various approaches for studying the effects of combination drug treatments. For the median effect model, a linear relationship between the logit transformed effect and the log-dose makes the model fitting straightforward and easy. However, when measuring cell growth, as in the experiments we analyzed, if the maximum drug effect reaches a plateau and does not kill all the cancer cells, even at the highest experimental doses, the median effect model (12) does not apply. We used the E_{max} model (19), which provides an adequate fit for most data. Parameter estimation under the E_{max} model requires the use of iterative procedures such as the nonlinear weighted least squares method, which can address the heteroscedascity problem. Model convergence is not guaranteed; whether or not the model converges depends on the data and the choice
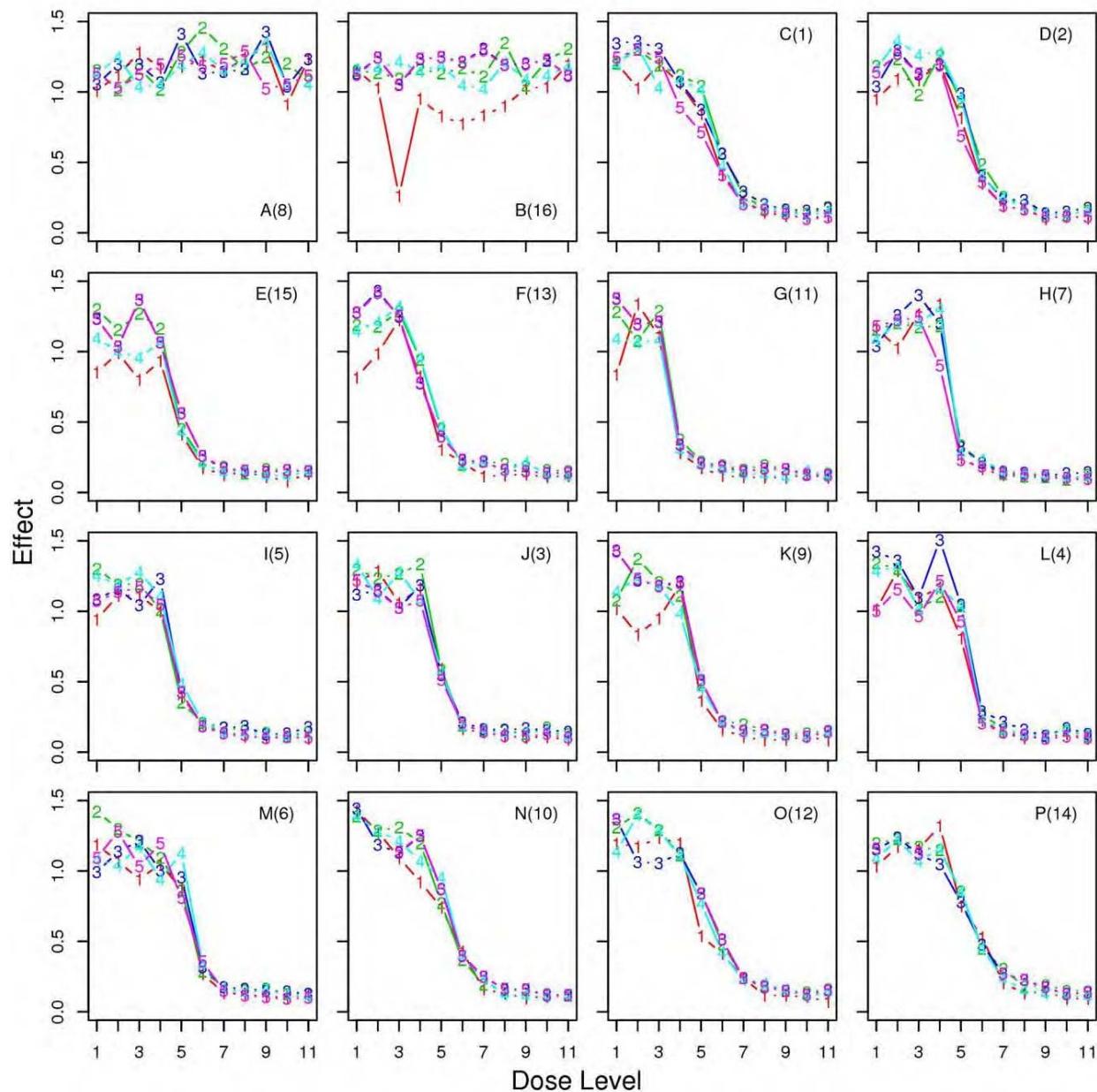
Figure 4. Distribution of the effect versus dose level for curves A through P for the experiment in a low folic acid medium.

of the initial values. We find that PROC NLIN in SAS provides a more comprehensive and robust environment for estimating parameters with nonlinear regression compared to the nls() function in S-PLUS/R. It can be useful to apply SAS first to estimate the parameters and then feed the results into S-PLUS/R for further data analysis and production of graphics. Unlike fitting the linearly-transformed median effect model via linear regression, for which a solution can always be found, fitting the E_{max} model via nonlinear regression may result in nonconvergence of the model in some cases. This nonconvergence may indicate aberrant conditions in the data such that the data do not provide adequate information

for model fitting. We had convergence problems with the curves G, H, and K in the low FA experiment. In these cases, there were insufficient data in the middle of the effect range; hence, the parameters could not be estimated reliably. We had to fix the *m* parameter before we could estimate the other two parameters. From the dose-response curves, we found that TMQ was more potent than AG, and that the drug combination was more potent in the low FA medium than in the high FA medium.

Upon construction of the marginal and combination dose-response curves, we applied the Loewe additivity model to compute the interaction index. We note
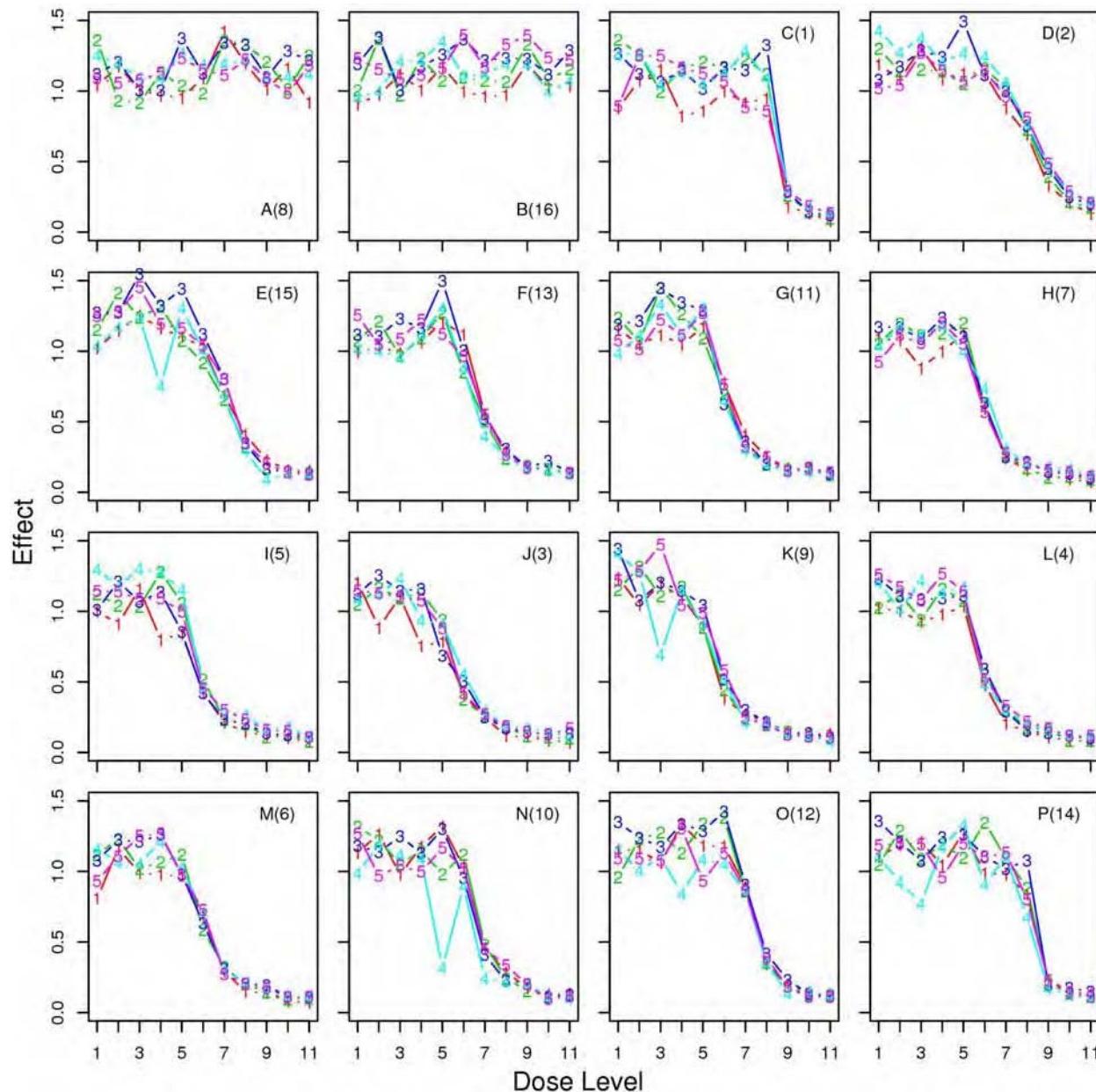
Figure 5. Distribution of the effect versus dose level for curves A through P for the experiment in a high folic acid medium.

that a definition of drug interaction such as the interaction index is model dependent. Additionally, no matter which model is used, based on the definition of the interaction index (7,8), the dose levels used in calculating the interaction index must be translated back to the original units of dose measurement. Under the given model, we found that the drug interaction between TMQ and AG was largely synergistic. Synergy was more clear and evident in the high FA experiment than in the low FA experiment. In addition, synergy was more likely to be observed when a small dose of the more potent drug (TMQ) was added to a large dose of the less potent drug (AG). When a large amount of a

more potent drug is present, adding the less potent drug does not show synergy because the effect is already largely achieved by the more potent drug. In addition, the interval estimation showed that the 95% confidence intervals were wider at the two extremes of the effect, which were closer to 1 or to $1-E_{max}$. This result is consistent with that of many regression settings in which estimation achieves higher precision in the center of the data distribution but lower precision at the extremes.

We have provided a simple, yet useful approach for analyzing drug interaction for combination studies. The interaction index for each fixed dose ratio
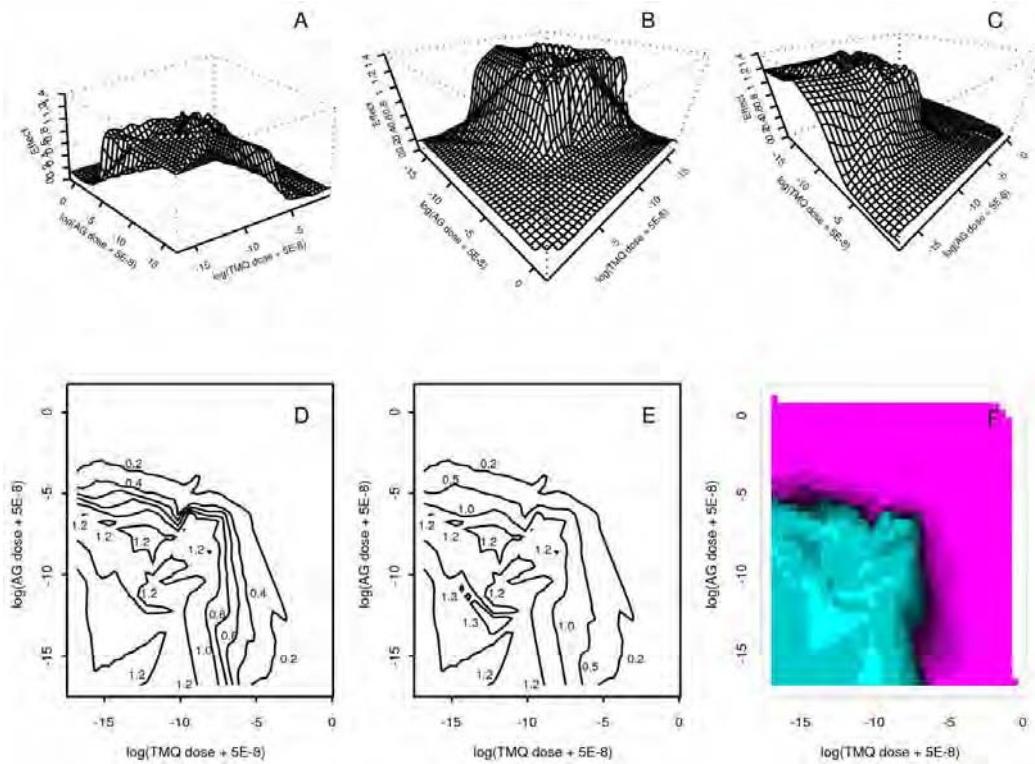
Figure 6. Perspective plots (A, B, C), contour plots (D, E), and image plot (F) for the effect versus logarithm transformed doses of trimetrexate and AG2034 for the experiment in a low folic acid medium.
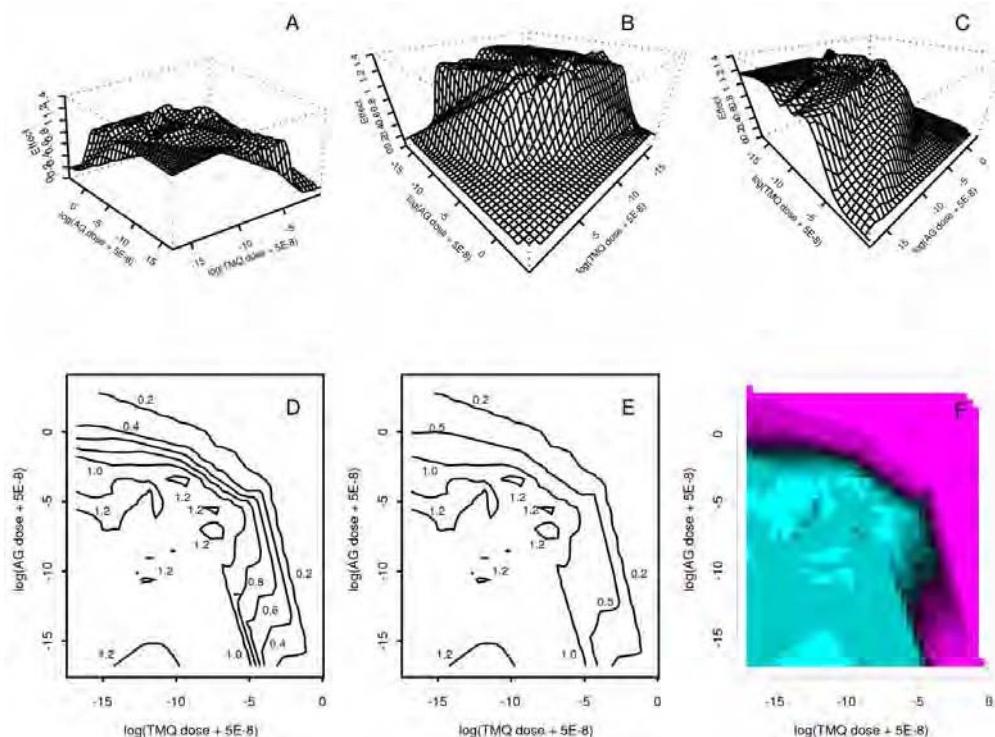
Figure 7. Perspective plots (A, B, C), contour plots (D, E), and image plot (F) for the effect versus logarithmically transformed doses of trimetrexate and AG2034 for the experiment in a high folic acid medium.
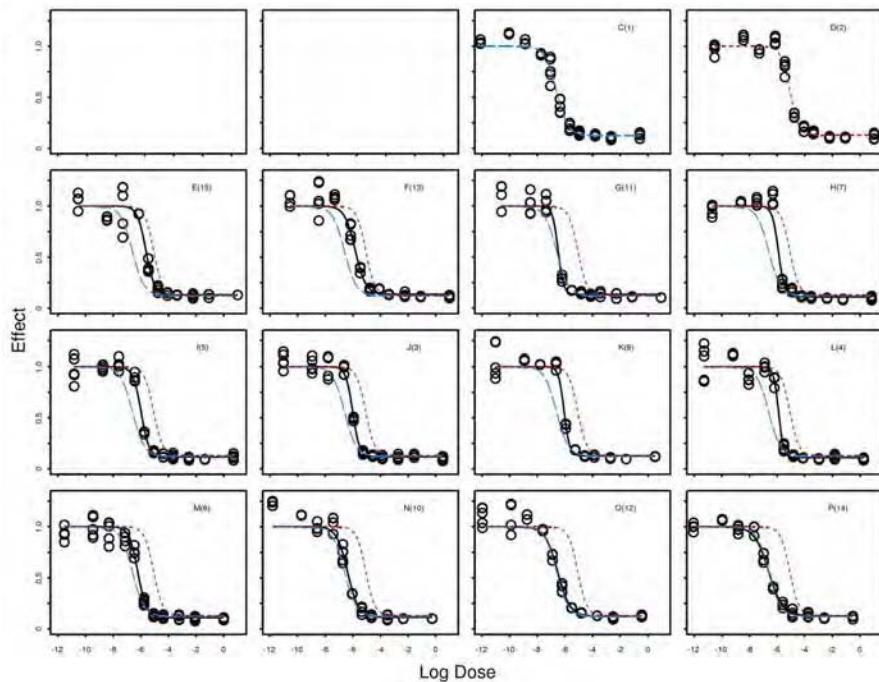
Figure 8. Effect versus logarithmically transformed dose plot for the combination study of trimetrexate and AG2034 in a low folic acid medium. Raw data are shown in open circles. Blue dashed line and red dotted line indicate the fitted marginal dose-response curves for trimetrexate and AG2034, respectively. Black solid line indicates the fitted dose-response curve for the combination of trimetrexate and AG2034.
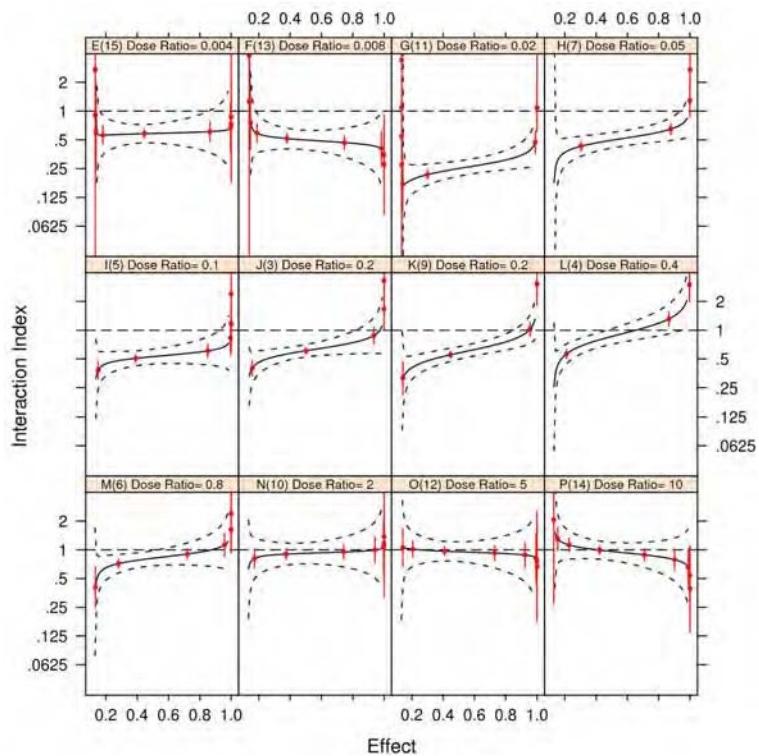
Figure 9. Trellis plot of the estimated interaction index (solid line) and its point-wise 95% confidence interval (red solid lines) and the 95% simultaneous confidence band (dashed lines) for the low folic acid experiment. Estimates at the design points where experiments were conducted are in red. The interaction index is plotted on the logarithmically transformed scale but labeled on the original scale.
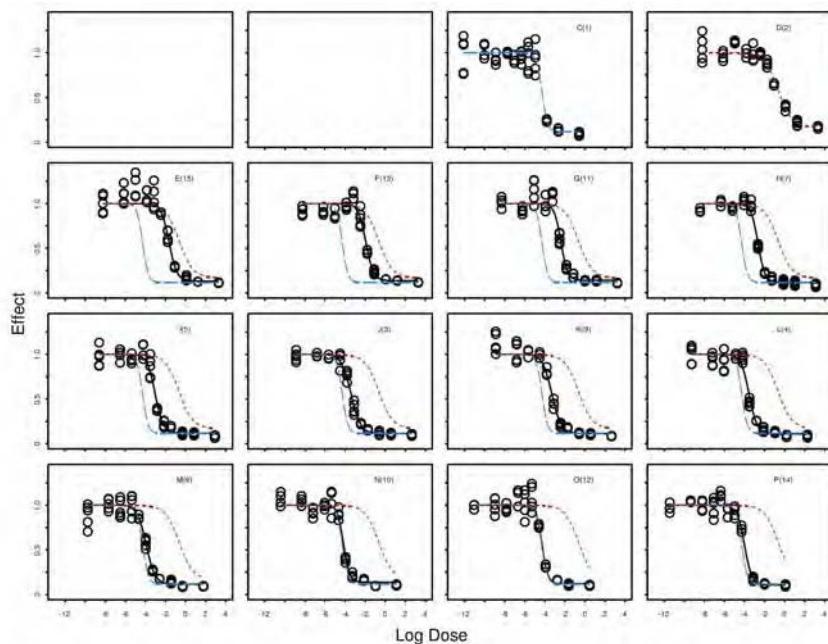
Figure 10. Effect versus logarithmically transformed dose plot for the combination study of trimetrexate and AG2034 in a high folic acid medium. Raw data are shown in open circles. Blue dashed line and red dotted line indicate the fitted marginal dose-response curves for trimetrexate and AG2034, respectively. Black solid line indicates the fitted dose-response curve for the combination of trimetrexate and AG2034.
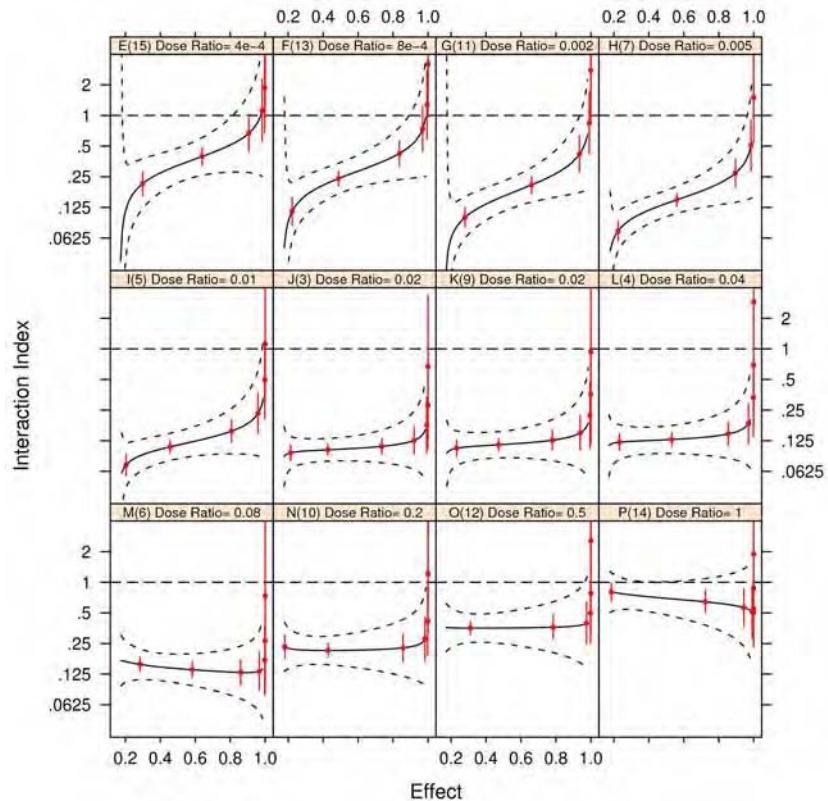
Figure 11. Trellis plot of the estimated interaction index (solid line) and its point-wise 95% confidence interval (red solid lines) and the 95% simultaneous confidence band (dashed lines) for the high folic acid experiment. Estimates at the design points where experiments were conducted are in red. The interaction index is plotted on the logarithmically transformed scale but labeled on the original scale.

is computed and then displayed together using a trellis plot. This method works well for the ray design. Other methods have been proposed to model the entire response surface using the parametric approach (27) or the semiparametric approach (28). The results from applying the semiparametric model are reported in a companion article (29).

9. ACKNOWLEDGEMENT

10. REFERENCES

1. R. Chris Rathbun, Staci M. Lockhart, Johnny R. Stephens: Current HIV treatment guidelines - an overview. Curr Pharm Des 12, 1045-1063 (2006)

2. Julian R. Molina, Alex A. Adjei, James R. Jett: Advances in chemotherapy of non-small cell lung cancer. Chest 130, 1211-1219 (2006)

3. Myaing M. Nyunt, Christopher V. Plowe: Pharmacologic advances in the global control and treatment of malaria: combination therapy and resistance. Clin Pharmacol and Ther 82, 601-605 (2007)

4. Kamalesh K. Sankhala, Kyriakos P. Papadopoulos: Future options for imatinib mesilate-resistant tumors. Expert Opinion Investig Drugs 16, 1549-1560 (2007)

5. Roberto Bianco, Vincenzo Damiano, Teresa Gelardi, Gennaro Daniele, Fortunato Ciardiello, Giampaolo Tortora: Rational combination of targeted therapies as a strategy to overcome the mechanisms of resistance to inhibitors of EGFR signaling. Curr Pharm Des 13, 3358-3367 (2007)

6. Sith Sathornsumetee, David A. Reardon, Annick Desjardins, Jennifer A. Quinn, James J. Vredenburgh, Jeremy N. Rich: Molecularly targeted therapy for malignant glioma. Cancer 110, 12-24 (2007)

7. Morris C. Berenbaum: What is synergy? Pharmacol Rev 41, 93-141 (1989)

8. William R. Greco, Gregory Bravo, John C. Parsons: The search of synergy: A critical review from a response surface perspective. Pharmacol Rev 47(2), 331-385 (1995)

9. Jurgen Suhnel: Parallel dose-response curves in combination experiments. Bull Math Biol 60, 197-213 (1998)

10. Ting-Chao Chou: Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. Pharmacol Rev 58, 621-681 (2006)

11. Ronald J. Tallarida: An overview of drug combination analysis with isobolograms. J Pharmacol Exp Ther 319, 1-7 (2006)

12. Ting-Chao Chou, Paul Talalay: Quantitative analysis of dose effect relationships: the combined effects of multiple drugs or enzyme inhibitors. Adv Enzyme Reg 22, 27-55 (1984)

13. Stella G. Machado, Grant A. Robinson: A direct, general approach based on isobolograms for assessing the joint action of drugs in pre-clinical experiments. Stat Med 13, 2289–2309 (1994)

14. Chris Gennings: On testing for drug/chemical interactions: definitions and inference. J Biopharm Stat 10, 457-467 (2000)

15. Kathryn S. Dawson, Walter H. Carter, Jr., Chris Gennings: A statistical test for detecting and characterizing departures from additivity in drug/chemical combinations. J Agric Biol Environ Stat 5, 342-359 (2000)

16. J. Jack Lee, Maiying Kong, Gregory D. Ayers, Reuben Lotan: Interaction index and different methods for determining drug interaction in combination therapy. J Biopharm Stats 17, 461-480 (2007)

17. J. Jack Lee, Maiying Kong: A confidence interval for interaction index for assessing multiple drug interaction. Stat Biopharm Res 1, 4-17, (2009)

18. Helene M. Faessel, Harry K. Slocum, Robert C. Jackson, Theodore J. Boritzki, Youcef M. Rustum, M. G. Nair, and William R. Greco: Super in vitro synergy between inhibitors of dihydrofolate reductase and inhibitors of other folate-requiring enzymes: The critical role of polyglutamylation. Cancer Res 58, 3036-3050 (1998)

19. Naitee Ting: Dose finding in drug development. Springer, NY, 127-145 (2006)

20. John M. Chambers, Trevor J. Hastie: Statistical models in S. Chapman & Hall/CRC Press, Boca Raton FL, 450 – 452 (1992)

21. Sylvie Huet, Anne Bouvier, Marie-Anne Poursat, Emmanuel Jolivet: Statistical tools for nonlinear regression: a practical cuide with S-PLUS and R examples, 2nd ed. Springer, NY (2003)

22. SAS/STAT 9.1 User's guide. SAS Institute, Cary, NC (2007)

23. Peter J. Bickel, Kjell A. Doksum: Mathematical statistics: basic ideas and selected topics, Vol I, 2nd ed. *Prentice Hall*, New Jersey pp. 306-314 (2000)

24. Christopher Cox, Guangpeng Ma: Asymptotic confidence bands for generalized nonlinear regression models. *Biometrics* 51, 142-150 (1995)

25. J. Jack Lee, Z. Nora Tu: A versatile one-dimensional distribution plot: the BLiP plot. *Amer Statistician* 51(4), 353-358 (1997)

26. Michael J. Crawley: The R book. *Wiley*, NY (2007)

27. Maiying Kong, J. Jack Lee: A general response surface model with varying relative potency for assessing drug interactions. *Biometrics* 62(4), 986-995 (2006)

28. Maiying Kong, J. Jack Lee: A semiparametric response surface model for assessing drug interactions. *Biometrics* 64, 396-405 (2008)

29. Maiying Kong, J. Jack Lee: Applying E_{max} model and bivariate thin plate splines to assess drug interactions. *Front Biosci* In Press

Abbreviations: AG: AG2034, an inhibitor of the enzyme glycinamide ribonucleotide formyltransferase; ED_{50}: dose required to produce 50% of the maximum effect; E_{max}: maximum effect attributed to the drug; FA: folic acid; II: interaction index; TMQ: Trimetrexate, a lipophilic inhibitor of the enzyme dihydrofolate reductase;

Key Words: additivity, antagonism, confidence interval estimation, E_{max} model, Loewe additivity model, nonlinear regression, synergy, trellis plot

Send correspondence to: J. Jack Lee, Department of Biostatistics, Unit 1411, The University of Texas M. D. Anderson Cancer Center, P. O. Box 301402, Houston, Texas 77230-1402, Tel: 713-794-4158, Fax: 713-563-4242, E-mail: jjlee@mdanderson.org

http://www.bioscience.org/current/vol2E.htm

The Interleukin-11 Receptor α as a Candidate Ligand-Directed Target in Osteosarcoma: Consistent Data from Cell Lines, Orthotopic Models, and Human Tumor Samples

Valerae O. Lewis,[1] Michael G. Ozawa,[2] Michael T. Deavers,[3] Guiying Wang,[1] Tomo Shintani,[4] Wadih Arap,[2] and Renata Pasqualini[2]

[1]Section of Orthopedic Oncology, Department of Surgical Oncology, [2]David H. Koch Center, and Departments of [3]Pathology and [4]Radiation Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas

Abstract

The interleukin-11 receptor α (IL-11Rα) is a functional target in bone metastasis. However, its role in primary bone tumors has not been established. As such, here, we evaluated IL-11Rα as a candidate target in primary and metastatic human osteosarcoma. First, in an orthotopic mouse model, we showed that IL-11Rα protein is markedly expressed in primary osseus and pulmonary metastatic osteosarcoma but absent from control normal tibia and lung. Moreover, systemic administration of an IL-11Rα–targeting phage displaying the cyclic nonapeptide CGRRAGGSC resulted in strong and selective accumulation of IL-11Rα–homing phage particles in the osteosarcoma but not in several control organs. Finally, IL-11Rα expression in a large panel of human primary and metastatic osteosarcoma samples was remarkably consistent with the observations in the orthotopic mouse model. These data establish IL-11Rα as a candidate target in human osteosarcoma and provide leads for the development of novel imaging and therapeutic agents for the management of this malignant tumor. [Cancer Res 2009;69(5):1995–9]

Introduction

Human osteosarcoma is the most common primary malignant tumor of bone (1). Although the introduction of modern chemotherapy has improved the 5-year survival to nearly 70%, hearing loss, cardiomyopathy, neuropathy, and renal failure are seriously debilitating and even fatal side effects of the systemic combination cytotoxic chemotherapy currently used (2, 3). Furthermore, tumors often respond to chemotherapy, but ultimately, many patients succumb to respiratory failure secondary to progression of pulmonary metastases. Investigators have evaluated many alternatives for selective therapy, including tyrosine kinase inhibitors, insulin-like growth factor-I inhibitors, radionucleotides, gene therapy, and monoclonal antibodies such as disialoganglioside GD2 and gp58 (refs. 4–9). However, these newer noncytotoxic treatments either lack systemic specificity or have limited efficacy against human sarcomas; thus, additional treatment options for this tumor is clearly an unmet need.

Molecular heterogeneity of the vascular endothelium enables selective targeting of agents to normal or diseased tissues (10, 11). Combinatorial approaches, such as *in vivo* phage display library selection, allow for interrogation of the endothelial cell surface in their native microenvironment and identification of targeting peptide motifs suitable for ligand-directed delivery (12). In previous work, we screened a phage display library in a human subject and reported a nonrandom peptide distribution (13–15). One of the selected ligand peptides contained the targeting motif Arg-Arg-Ala-Gly-Gly-Ser (RRAGGS, single-letter code) and exhibited similarity to human interleukin-11 (IL-11). This IL-11 mimic peptide bound the IL-11 receptor α (IL-11Rα), through a previously unrecognized binding site (16), and localized to the vasculature of human prostate cancer (13). Moreover, evaluation of IL-11Rα expression in an expanded set of clinically annotated prostate cancer samples has also shown a strong increase and gradual epithelial expression of IL-11Rα with concomitant pathologic progression to bone metastases (17). On an unrelated line of evidence, IL-11/IL-11Rα binding and downstream signaling via signal transduction and activator of transcription 3 activation has been proposed as a leading molecular pathway in metastasis (18).

In the present study, given the high expression of IL-11Rα protein and the proposed role of IL-11/IL-11Rα signaling in bone metastasis, we hypothesized that IL-11Rα similarly serves as a candidate target for primary and metastatic osteosarcoma. Using morphologic and functional analyses, we first examined a panel of mouse and human osteosarcoma–derived cell lines to establish the presence of a functional IL-11Rα protein. In animal models of osteosarcoma, we then showed that IL-11Rα within the bone microenvironment is accessible to a circulating particle displaying a mimic of the native ligand, IL-11, and strongly accumulates within the tumor. Finally, we showed that human primary osteosarcoma and pulmonary metastases express IL-11Rα, both in tumor cells and activated tumor blood vessels, while sparing normal bone marrow and lung. Together, these data indicate for the first time that IL-11Rα is a candidate target in human osteosarcoma and may serve as a target for ligand-directed delivery of agents against this disease.

Materials and Methods

Cell culture. Osteosarcoma cell lines derived from human (KRIB and OS187) and mouse (Dunn-LM, K7M3) were grown in DMEM containing 10% FCS, sodium pyruvate, streptomycin, nonessential amino acids, and multivitamins (4, 19).

Flow cytometric analysis of IL-11Rα expression. Osteosarcoma cells were plated at 8×10^5 cells. After 24 h, cells were trypsinized, suspended in a fluorescence-activated cell sorting buffer consisting of PBS, 2% FCS, and 0.1% NaN₃, and incubated with either 1 mg/mL phycoerythrin

(PE)-conjugated hamster anti-mouse IL-11 monoclonal antibody or isotype-matched, PE-conjugated control hamster IgG antibody (Pharmingen) for 30 min at 4°C. Samples were washed and analyzed by flow cytometry (FACScan; Becton Dickinson).

Cell internalization assay. Internalization assays were performed as described (17). In brief, osteosarcoma cells were plated overnight on 6-well polystyrene plates to a concentration of 5×10^4 cells/2 mL medium. Cells were blocked with 500 μL of DMEM containing 30% FCS for 60 min at 37°C then incubated with 10^9 transducing units (TU) of phage in DMEM containing 2% FCS for 4 h at 37°C. Cells were sequentially washed with PBS containing 10% bovine serum albumin (BSA), glycine buffer, PBS, and then fixed with PBS containing 4% paraformaldehyde (PFA). Cells were permeabilized with PBS containing 0.2% Triton X-100 (5 min at room temperature; RT), extensively washed with PBS, and blocked with PBS containing 5% normal serum and 1% BSA for 2 h at RT. The cells were incubated with a rabbit antiphage antibody (Sigma; 1:500 dilution) in PBS containing 1% normal serum (2 h at RT) and washed with PBS. Primary antibodies were detected with a Cy3-conjugated anti-rabbit IgG (Jackson ImmunoResearch; 1:300 dilution) in PBS (1 h at RT), rinsed with PBS (1 min at RT), fixed with PFA (15 min at RT), and mounted with Vectashield (Vector Laboratories).

Intratibial implantation of osteosarcoma cells. Male athymic nu/nu (nude) mice were purchased from the animal production area of the National Cancer Institute-Frederick Cancer Research Facility and maintained in a pathogen-free barrier animal facility approved by the American Association for Accreditation of Laboratory Animal Care.

For intratibial administration, cultured osteosarcoma cells at 80% confluence were briefly washed with PBS and then detached with 0.25% trypsin and 0.02% EDTA. Cell detachment was stopped after 1 min with DMEM containing 10% FCS. Cells were washed once in serum-free medium and resuspended in Hank's buffered salt solution. Suspensions containing >90% viable (trypan blue–excluding) cells were used for administration. Cells were injected (5×10^4 cells per mouse) into the right tibia of 5-wk-old nude mice anesthetized with Nembutal (50 mg/kg i.p.). Tumor growth was monitored with a Faxitron MX-20 X-ray unit, and images were digitally captured. Mice were killed at 1, 3, 5, 7, and 8 wk postinjection. Tibias

(experimental and contralateral) and visceral organs were collected and processed. Histopathology was used to evaluate the presence of tibial and/or pulmonary tumors.

Immunohistochemical analysis. Mouse osteosarcoma samples were stained within 2 wk of sectioning. Four-micrometer sections were subjected to antigen-retrieval by heat with EDTA (pH 8.0; Zymed) followed by biotin and protein blocking (DAKO). Cell expression of IL-11Rα protein was evaluated with a rabbit anti–IL-11Rα antibody (C20; Santa Cruz Biotechnology) that cross-reacts (human and mouse) with the receptor. Primary antibodies were diluted at 1:15 and incubated for 45 min followed by development with the LSAB+ kit (DAKO).

Dual staining for IL-11Rα and CD31. Tumor specimens were collected and analyzed for IL-11Rα and CD31 expression as described (20, 21). Briefly, frozen tissue cryostat sections were fixed in a 50:50 solution of alcohol and acetone for 15 min, then washed with PBS. Nonspecific proteins were blocked by incubation in 5% normal horse serum plus 1% normal goat serum in PBS for 20 min. Sections were incubated with an anti-CD31 primary antibody followed by a secondary antibody conjugated with Alexa 594 (Molecular Probes). Sections were blocked again with PBS containing 4% gelatin (Electron Microscopy Sciences) and washed with PBS. Protein expression of IL-11Rα was detected with an anti–IL-11Rα rabbit antibody (Santa Cruz), followed by a goat anti-rabbit antibody conjugated with Alexa 488 (Molecular Probes). Slides were counterstained with Hoechst 3222 and visualized under standard fluorescence microscopy.

Tumor targeting. Targeted phage experiments *in vivo* were performed as described (22–24). Briefly, male athymic nude mice bearing intratibial tumors were anesthetized and injected i.v. (tail vein) with 10^9 TU of CGRRAGGSC-displaying phage, RGD-4C phage (positive control), or an fd insertless phage (negative control). After 24 h, mice were systemically perfused through the heart with 20 mL of PFA. Tumor and control organs were removed and fixed in formalin. Tibial tumors underwent decalcification. Tissue samples were paraffin-embedded and sectioned into 4-μm specimens for staining (23).

IL-11Rα in human primary tumors and lung metastases. After Institutional Review Board approval, the osteosarcoma database and hospital tumor registry at the University of Texas M. D. Anderson Cancer
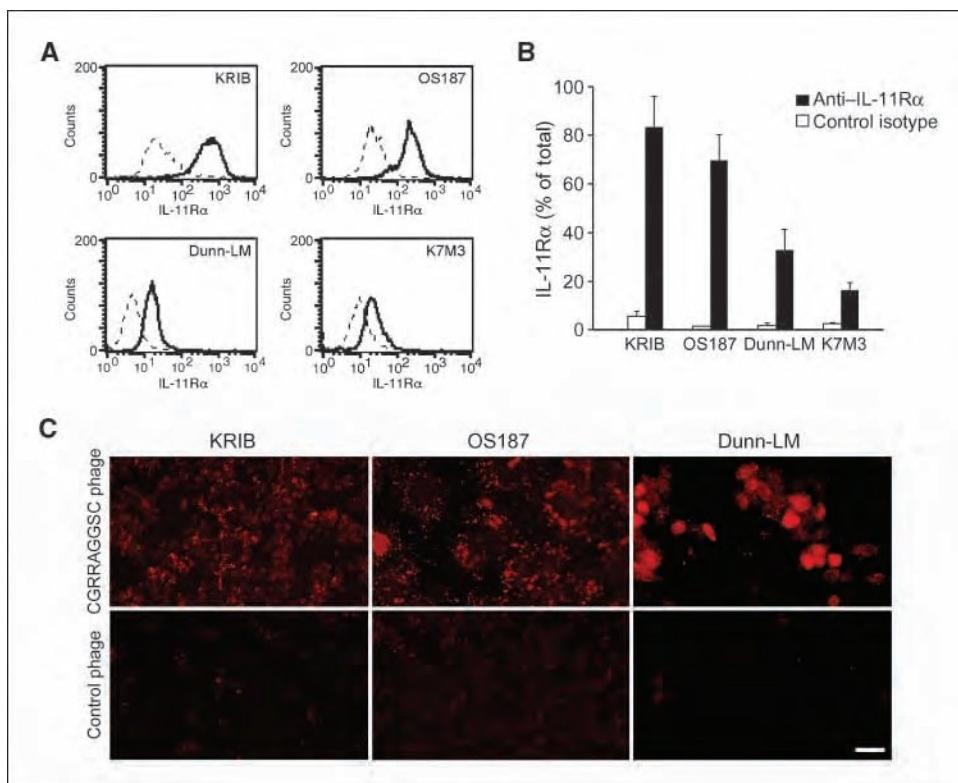
Figure 1. IL-11Rα protein expression in osteosarcoma cells. *A,* cells were labeled with either PE-conjugated polyclonal anti–IL-11Rα antibody (*solid bold line*) or PE-conjugated IgG isotype control (*dashed line*) and measured by flow cytometry. *B,* expression of IL-11Rα was statistically analyzed (*Sigma Plot*). The data represent mean values of three independent experiments (*t* test, *P* < 0.05). *C,* IL-11Rα–targeted peptides mediate receptor-ligand internalization. CGRRAGGSC-displaying phage or control insertless phage were incubated with osteosarcoma cell lines for 4 h at 37°C to allow targeted phage internalization. Internalized phage clones (*red dots*) were detected with an antiphage antibody after cell permeabilization. *Scale bar,* 30 μm.
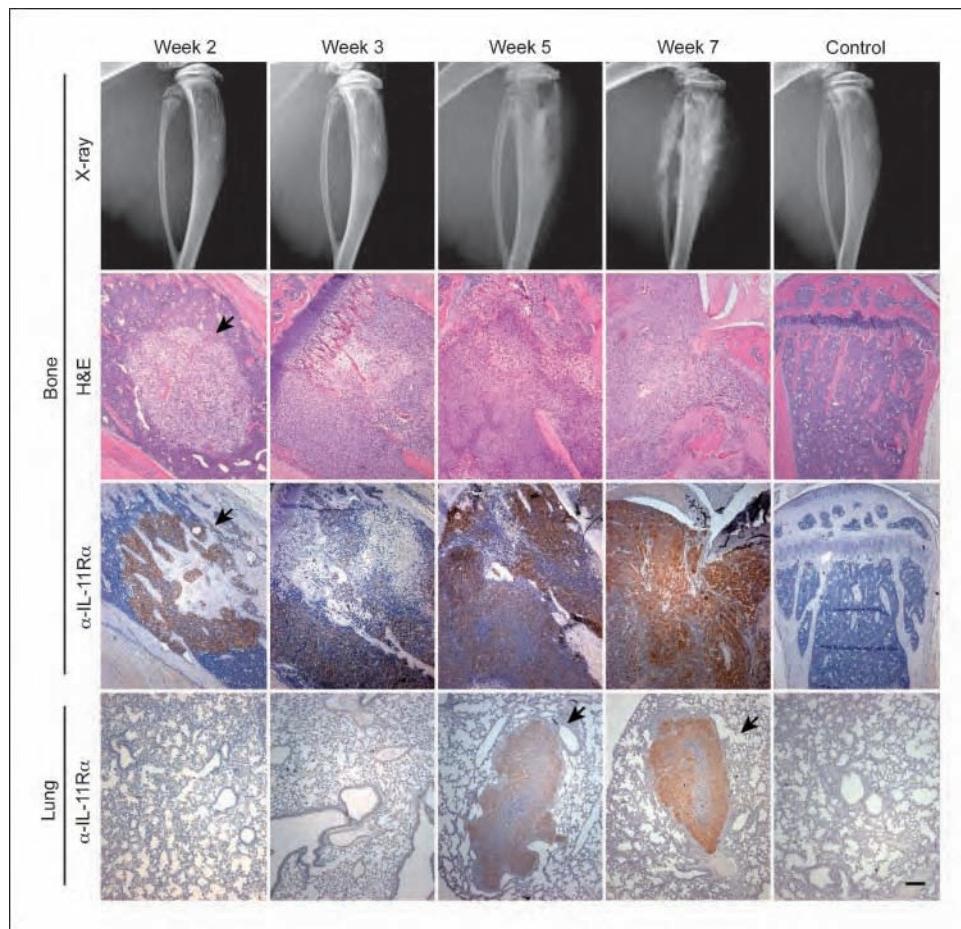
Figure 2. Serial radiographs and immunohistochemical analysis of IL-11Rα in tibial osteosarcoma and lung metastases. Radiograph of tibial osteosarcoma reproducibly shows progression of the lesion over time. A typical experiment is shown. At week 2, a cortically confined lesion is seen (*arrow*). At week 3, posterior breakthrough is noted. Week 5 shows longitudinal extension of the lesion with progressive bony destruction and posterior and anterior cortical breakthrough. Week 7 illustrates bony destruction and the classic sunburst pattern of periosteal reaction. In representative images of the tumors at 5 and 7 wk, the osteosarcoma have broken out anteriorly and posteriorly of the tibia. Control limbs show no visible lesion. Corresponding H&E staining and IL-11Rα expression of the tibial osteosarcoma lesions shows consistent expression of IL-11Rα with no expression in control normal bone (magnification, ×200). Lung sections stained for IL-11Rα expression show no expression in the control normal lung parenchyma but metastatic lesions express the receptor markedly. Pulmonary metastases were first detected at week 5 (*arrows*). *Scale bar,* 100 μm.

Center were queried from 2002 to 2007 for human osteosarcomas. Samples with <60% necrosis were then identified. From this specific query, primary bone samples (n = 30) and lung metastases (n = 19) were examined for the expression of the IL-11Rα. Slides with viable tumor were stained as detailed above. The specimens were graded for intensity of staining on a scale of 1 to 3 and for distribution on scale of 1 to 4 (Supplementary Fig. S1).

Results and Discussion

Osteosarcoma cells express IL-11Rα. To determine whether IL-11Rα would serve as a molecular target for osteosarcoma, we first evaluated the cell surface expression of IL-11Rα in several osteosarcoma cell lines. An anti–IL-11Rα antibody recognizing both human and mouse IL-11Rα showed positive reactivity of KRIB and OS187 (human) as well as Dunn-LM and K7M3 (mouse) cells (Fig. 1*A*). The data represent mean values of three independent experiments and are quantified (Fig. 1*B*). To further assess whether IL-11Rα is functionally active in osteosarcoma, we investigated the ability of KRIB, OS187, and Dunn-LM osteosarcoma cells to bind and internalize an IL-11 mimic peptide (17). Immunofluorescence analysis revealed that such CGRRAGGSC-displaying phage specifically bound and internalized into osteosarcoma cells, whereas an insertless negative control phage could not be detected under identical experimental conditions (Fig. 1*C*). These data support the expression of IL-11Rα on osteosarcoma cells and present a viable candidate target for ligand-directed delivery to tumor cells.

Systemic targeting of an orthotopic osteosarcoma model. *In vivo* models of osteosarcoma that spontaneously metastasize to

the lung are robust and reproducible systems on which to not only study the mechanism of tumor growth and pulmonary metastases but also to assess the efficacy of potential antitumor and anti-metastatic agents. There are several osteosarcoma cell lines that when injected orthotopically into the tibia develop evidence of clinical and radiographic osteosarcoma lesions (4, 19, 25). In addition, as in human osteosarcomas, these lesions can metastasize from the primary site into pulmonary metastases. The similarity of these rodent models to the natural history of osteosarcoma provide a valuable tool to rapidly translate newly identified molecular markers for clinical applications.

By using a representative panel of osteosarcoma cell lines (KRIB, OS187, Dunn-LM, and K7M3), we generated orthotopic, metastatic mouse models of osteosarcoma (4, 19). We established a time course of tumor development and monitored its burden by X-ray radiography. The histopathologic and experimental findings with each cell line were reproducibly similar. Therefore, we focused study on the KRIB cell–derived orthotopic model and studied the protein expression of IL-11Rα *in vivo* by immunohistochemistry in formalin-fixed, paraffin-embedded tissue samples at 2, 3, 5, and 7 weeks postinjection (Fig. 2). Beginning at an early time point (3 weeks), IL-11Rα staining was strongly localized to the intratibial lesion then limited to the periphery of the lesions observed at later time point s (6–8 weeks). This change in staining pattern suggested the IL-11Rα staining was most prominent in viable tumor areas, as the central areas of the tumors became necrotic. IL-11Rα expression in control normal bone (contralateral tibia) was barely
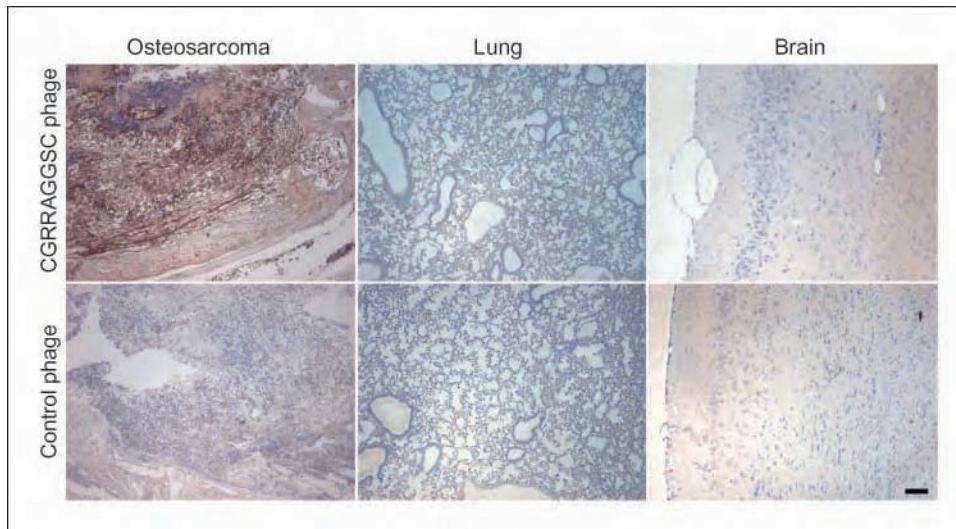
Figure 3. Systemic targeting of osteosarcoma by targeted phage. Immunohistochemical analysis of phage accumulation after i.v. administration of either CGRRAGGSC-displaying phage (*top row*) or insertless fd-tet negative control phage (*bottom row*) into mice bearing orthotopic osteosarcoma showed strong phage staining within the tumor, whereas the negative control was barely detectable above background. *Scale bar,* 100 μm.

detectable. Furthermore, at later time points, IL-11Rα was strongly expressed in pulmonary lesions with no expression detected in the control normal lung parenchyma (Fig. 2).

Having shown the presence of IL-11Rα in orthotopic models of osteosarcoma, we next sought to evaluate systemic targeting of IL-11 mimic–displaying phage particles to the osteosarcoma lesions. We i.v. administered either the CGRRAGGSC-displaying phage or the insertless negative control phage to mice bearing intratibial osteosarcomas. After 24 hours of circulation, the CGRRAGGSC-displaying phage showed strong staining within the tumors with little to no accumulation detected in several control organs by immunohistochemistry (Fig. 3, *top row*). In contrast, the insertless negative control phage were barely detectable in tumors (Fig. 3, *bottom row*). As previously well established (26, 27), phage were identified in the spleen and liver, which are part of the reticuloendothelial system that nonspecifically clear phage
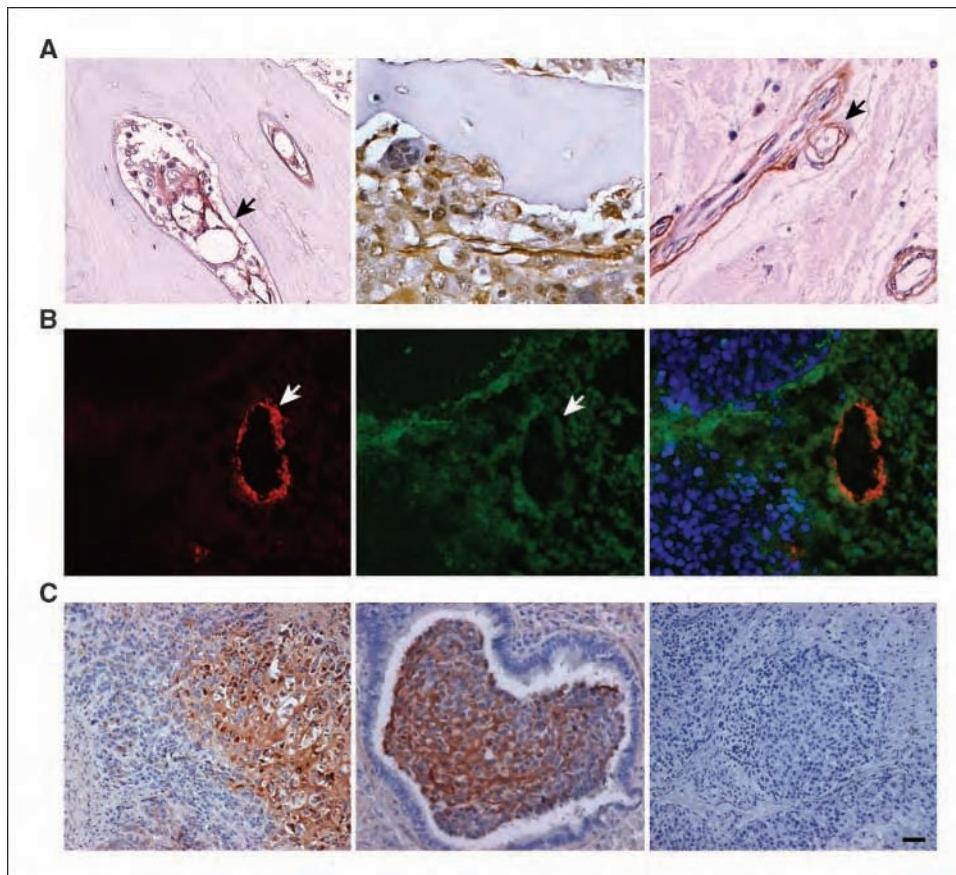
Figure 4. Immunohistochemical staining of IL-11Rα protein in human osteosarcoma biopsy samples and lung metastases. *A,* osteosarcoma cells present in Haversian canals (*left*) and the endothelial lining of small diameter blood vessels (*middle* and *right*) stain positively for IL-11Rα (*arrows*). No staining was observed in the bone. *B,* immunofluorescence detection of IL-11Rα in human osteosarcoma tumor vasculature showed localization of CD31 (*left*; *arrows*), IL-11Rα (*middle*; *arrows*), and colocalization (*right*) in human osteosarcoma. *C,* tumor cells positively express IL-11Rα, whereas the normal lung parenchyma was negative (*left*). Intense staining of osteosarcoma cells was also observed in a bronchiole of the lung (*middle*) with control staining appearing negative (right). *Scale bar,* 80 μm.

independent of the displayed peptide. This selective targeting of the IL-11 mimic phage upon i.v. administration indicates not only that IL-11Rα is accessible to circulating agents but also suggests that therapeutic agents with this ligand-directed system may yield improved targeted agents.

IL-11Rα protein is markedly expressed in primary and metastatic human osteosarcoma. To evaluate whether expression of IL-11Rα in the murine tumor models translates to human osteosarcoma, we first identified a large panel of human tumors in which the samples had <60% necrosis on pathology (*n* = 30). From this set, we performed immunohistochemical analysis of IL-11Rα expression and scored the staining intensity and distribution (Supplementary Fig. S1; Fig. 4*A*). Moderate-to-high intensity staining of the tumor cells was noted in all primary osteosarcoma samples, with an average score of 2.49 and was tightly distributed with an average score of 2.37 (Supplementary Fig. S1; Fig. 4*A*). Endothelial expression of IL-11Rα within the tumors was seen in >50% of tumor blood vessels and again showed moderate to high intensity staining (average, 2.53; Supplementary Fig. S1; Fig. 4*B*). To confirm endothelial cell localization, we costained IL-11Rα with the endothelial cell marker CD31 (Fig. 4*B*). Of note, only the small caliber blood vessels within the tumor positively expressed the receptor, whereas large tumor blood vessels did not express IL-11Rα at detectable levels; the reason for this observation remains unclear.

We next set out to evaluate whether IL-11Rα expression was maintained in lung metastatic tumors (*n* = 19). All pulmonary metastases were positive for IL-11Rα, revealing a high intensity of staining with an average of 3.0 and a moderate distribution of 2.84 (Supplementary Fig. S1; Fig. 4*C*). The control normal lung parenchyma was negative for IL-11Rα protein expression. Taken together, therapeutic targeting of IL-11Rα could potentially act as an antitumor, antiangiogenesis, and antimetastatic agent for the management of human osteosarcoma.

In summary, treatment options and survival outcomes of patients with osteosarcoma has all but plateaued over the past 20 year. The discovery of functional ligand receptor systems is a critical first step in the development of new targeted agents. Our histologic and functional findings establish that the IL-11/IL-11Rα system acts as a bona fide ligand receptor pair in osteosarcoma. Moreover, the vascular expression, systemic targeting, and human translation further support the candidacy of IL-11/IL-11Rα–based therapies or imaging agents for osteosarcoma patients.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

References

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2006. CA Cancer J Clin 2006;56:106–30.
2. Benjamin R, Chawla S, Murray J, et al. Preoperative chemotherapy for osteosarcoma: a treatment approach facilitating limb salvage with major prognostic indications. In: Jones S, Salmon S, editors. Adjuvant therapy of cancer IV. Philadelphia: Grune and Strattonl 1984. p. 601–10.
3. Hudson M, Jaffe MR, Jaffe N, et al. Pediatric osteosarcoma: therapeutic strategies, results, and prognostic factors derived from a 10-year experience. J Clin Oncol 1990;8:1988–97.
4. McGary EC, Weber K, Mills L, et al. Inhibition of platelet-derived growth factor-mediated proliferation of osteosarcoma cells by the novel tyrosine kinase inhibitor STI571. Clin Cancer Res 2002;8:3584–91.
5. Raile K, Hoflich A, Kessler U, et al. Human osteosarcoma (U-2 OS) cells express both insulin-like growth factor-I (IGF-I) receptors and insulin-like growth factor-II/mannose-6-phosphate (IGF-II/M6P) receptors and synthesize IGF-II: autocrine growth stimulation by IGF-II via the IGF-I receptor. J Cell Physiol 1994;159:531–41.
6. Mansky PJ, Liewehr DJ, Steinberg SM, et al. Treatment of metastatic osteosarcoma with the somatostatin analog OncoLar: significant reduction of insulin-like growth factor-I serum levels. J Pediatr Hematol Oncol 2002;24:440–6.
7. Anderson PM, Wiseman GA, Dispenzieri A, et al. High-dose samarium-153 ethylene diamine tetramethylene phophonate: low toxicity of skeletal irradiation in patients with osteosarcoma and bone metastases. J Clin Oncol 2002;20:189–96.
8. Heiner JP, Miraldi F, Kallick S, et al. Localization of GD2-specific monoclonal antibody 3F8 in human osteosarcoma. Cancer Res 1987;47:5377–81.
9. Modak S, Kramer K, Gultekin SH, Guo HF, Cheung NK. Monoclonal antibody 8H9 targets a novel cell surface antigen expressed by a wide spectrum of human solid tumors. Cancer Res 2001;61:4048–54.
10. Aird WC. Phenotypic heterogeneity of the endothelium: I. structure, function, and mechanisms. Circ Res 2007;100:158–73.
11. Aird WC. Phenotypic heterogeneity of the endothelium: II. Representative vascular beds. Circ Res 2007;100:174–90.
12. Ozawa MG, Zurita AJ, Dias-Neto E, et al. Beyond receptor expression levels: the relevance of target accessibility in ligand-directed pharmacodelivery systems. Trends Cardiovasc Med 2008;18:126–32.
13. Arap W, Kolonin MG, Trepel M, et al. Steps toward mapping the human vasculature by phage display. Nat Med 2002;8:121–7.
14. Pentz RD, Flamm AL, Pasqualini R, Logothetis CJ, Arap W. Revisiting ethical guidelines for research with terminal wean and brain-dead participants. Hastings Cent Rep 2003;33:20–6.
15. Pentz RD, Cohen CB, Wicclair M, et al. Ethics guidelines for research with the recently dead. Nat Med 2005;11:1145–9.
16. Cardó-Vila M, Zurita AJ, Giordano RJ, et al. A ligand peptide motif selected from a cancer patient is a receptor-interacting site within human interleukin-11. PLoS ONE 2008;3:e3452.
17. Zurita AJ, Troncoso P, Cardó-Vila M, et al. Combinatorial screenings in patients: the interleukin-11 receptor α as a candidate target in the progression of human prostate cancer. Cancer Res 2004;64:435–9.
18. Kang Y, Siegel PM, Shu W, et al. A multigenic program mediating breast cancer metastasis to bone. Cancer Cell 2003;3:537–49.
19. Berlin O, Samid D, Donthineni-Rao R, Akeson W, Amiel D, Woods VL, Jr. Development of a novel spontaneous metastasis model of human osteosarcoma transplanted orthotopically into bone of athymic mice. Cancer Res 1993;53:4890–5.
20. Hanavadi S, Martin TA, Watkins G, Mansel RE, Jiang WG. Expression of interleukin 11 and its receptor and their prognostic value in human breast cancer. Ann Surg Oncol 2006;13:802–8.
21. Mohamedali KA, Kedar D, Sweeney P, et al. The vascular-targeting fusion toxin VEGF121/rGel inhibits the growth of orthotopic human bladder carcinoma tumors. Neoplasia 2008;7:912–20.
22. Arap W, Pasqualini R, Ruoslahti E. Cancer treatment by targeted drug delivery to tumor vasculature in a mouse model. Science 1998;279:377–80.
23. Kolonin MG, Saha PK, Chan L, Pasqualini R, Arap W. Reversal of obesity by targeted ablation of adipose tissue. Nat Med 2004;10:625–32.
24. Arap MA, Lahdenranta J, Mintz PJ, et al. Cell surface expression of the stress response chaperone GRP78 enables tumor targeting by circulating ligands. Cancer Cell 2004;6:275–84.
25. Luu HH, Kang Q, Park JK, et al. An orthotopic model of human osteosarcoma growth and spontaneous pulmonary metastasis. Clin Exp Metastasis 2005;22:319–29.
26. Rajotte D, Arap W, Hagedorn E, Koivunen E, Pasqualini R, Ruoslahti E. Molecular heterogeneity of the vascular endothelium revealed by *in vivo* phage display. J Clin Invest 1998;102:430–7.
27. Pasqualini R, Koivunen E, Ruoslahti E. Alpha v integrins as receptors for tumor targeting by circulating ligands. Nat Biotechnol 1997;15:542–6.

Immunohistochemical Expression of Estrogen and Progesterone Receptors Identifies a Subset of NSCLCs and Correlates with *EGFR* Mutation

Maria G. Raso,[1] Carmen Behrens,[2] Matthew H. Herynk,[2] Suyu Liu,[3] Ludmila Prudkin,[1] Natalie C. Ozburn,[2] Denise M. Woods,[1] Ximing Tang,[2] Reza J. Mehran,[4] Cesar Moran,[1,2] J. Jack Lee,[3] and Ignacio I. Wistuba[1,2]

Abstract

Purpose: To determine the frequency of estrogen receptor α and β and progesterone receptor protein immunohistochemical expression in a large set of non–small cell lung carcinoma (NSCLC) specimens and to compare our results with those for some of the same antibodies that have provided inconsistent results in previously published reports.

Experimental Design: Using multiple antibodies, we investigated the immunohistochemical expression of estrogen receptors α and β and progesterone receptor in 317 NSCLCs placed in tissue microarrays and correlated their expression with patients' clinicopathologic characteristics and in adenocarcinomas with *EGFR* mutation status.

Results: Estrogen receptors α and β were detected in the nucleus and cytoplasm of NSCLC cells; however, the frequency of expression (nucleus, 5-36% for α and 42-56% for β; cytoplasm: <1-42% for α and 20-98% for β) varied among the different antibodies tested. Progesterone receptor was expressed in the nuclei of malignant cells in 63% of the tumors. Estrogen receptor α nuclear expression significantly correlated with adenocarcinoma histology, female gender, and history of never smoking (*P* = 0.0048 to <0.0001). In NSCLC, higher cytoplasmic estrogen receptor α expression significantly correlated with worse recurrence-free survival (hazard ratio, 1.77; 95% confidence interval, 1.12, 2.82; *P* = 0.015) in multivariate analysis. In adenocarcinomas, estrogen receptor α expression correlated with *EGFR* mutation (*P* = 0.0029 to <0.0001). Estrogen receptor β and progesterone receptor but not estrogen receptor α expressed in the normal epithelium adjacent to lung adenocarcinomas.

Conclusions: Estrogen receptor α and β expression distinguishes a subset of NSCLC that has defined clinicopathologic and genetic features. In lung adenocarcinoma, estrogen receptor α expression correlates with *EGFR* mutations. (Clin Cancer Res 2009;15(17): 5359–68)

Lung cancer is the most common cause of cancer mortality worldwide, with >1 million deaths each year (1). Lung cancer includes several histologic types, the most frequently occurring of which are two types of non–small cell lung carcinoma (NSCLC): adenocarcinoma and squamous cell carcinoma (2). During the last two decades, mortality rates associated with cancer have continued to decrease across all major sites in men and women; however, the rates for lung cancer in females have continued to increase (3, 4). Despite global statistics estimating that 15% of lung cancer in men and 53% in women are not attributable to smoking (1), smoking remains the primary risk factor for lung cancer. The higher proportion of lung cancer in females who have never smoked compared with males who have never smoked suggests a possible role for gender-dependent hormones in the development of lung cancer (5).

Estrogen receptors α and β are expressed in normal lung tissue and in lung tumors in men and women (6), yet the data are inconsistent about whether estrogen receptor expression is

Translational Relevance

A better understanding of the signaling pathways that lead to tumor growth may help in the development of new and more effective strategies for targeted chemoprevention and treatment of lung cancer. Our finding of frequent overexpression of estrogen receptor and progesterone receptor in non–small cell lung carcinoma suggests that the activation of these pathways is an attractive novel target for lung cancer chemopreventive and therapeutic strategies. The correlation between estrogen receptor and *EGFR* mutation in lung adenocarcinoma suggests that it might be important to target both pathways simultaneously for lung cancer therapy.

gender biased (6–9) or associated with NSCLC overall survival (9–11). The data reported on the immunohistochemical expression for both estrogen receptors in NSCLC remain controversial. Estrogen receptor α has been reported to be expressed in the nucleus (0-45%) and cytoplasm (0-73%) of malignant lung cancer cells in the cases examined (9, 10, 12, 13). The percentages for estrogen receptor β are more consistent, with 46% to 60% of NSCLC cases showing only nuclear expression (9–14). Similarly, two reports suggested that progesterone receptor is frequently (47%) expressed in NSCLC tumor cells, and this expression correlated with better patient outcome (12, 15).

Several *in vitro* and *in vivo* studies have provided evidence supporting a biological role for estrogens in lung carcinogenesis by direct promotion of cell-proliferation estrogens stimulate the proliferation of NSCLC cells through estrogen receptor–mediated signaling, whereas antiestrogens inhibit the growth of NSCLC cells (6, 7, 13, 16, 17). Estrogen can directly stimulate the transcription of estrogen-responsive genes in the nucleus of lung cells and can also transactivate growth factor–signaling pathways, the epidermal growth factor receptor (EGFR) pathway in particular (13, 18). In estrogen stimulation of lung cancer cells, EGFR ligands are rapidly released, activating the EGFR and mitogen-activated protein kinase 1 growth pathways (19). Activation of the EGFR pathway seems to play an important role in the pathogenesis and progression of NSCLC (20). In lung cancer cells, the constitutive activation of EGFR is achieved by several mechanisms, including increased production of ligands, increased levels of the receptor, and mutation of the *EGFR* tyrosine kinase domain (20–22). Of interest, EGFR protein expression is down-regulated in response to estrogens and up-regulated in response to antiestrogens, suggesting that a reciprocal control mechanism exists between the EGFR and estrogen receptor pathways (19).

The purpose of the current study was to determine the frequency of estrogen receptor α and β and progesterone receptor protein immunohistochemical expression in a large set of NSCLCs placed in tissue microarray specimens and to compare our results with those for some of the same antibodies that have provided inconsistent results in previously published reports (9–14). In addition, the receptor-expression results were correlated with patients' clinicopathologic features, including NSCLC histology, gender, smoking history, and patient outcome, and in adenocarcinoma with tumors' *EGFR* activating

mutation status. Finally, to understand estrogen receptor α and β and progesterone receptor protein expression in the early pathogenesis of lung cancer, we investigated the characteristics of estrogen receptor α and β and progesterone receptor protein expression in the nonmalignant respiratory epithelium adjacent to tumors taken from a subset of our retrospectively reviewed lung adenocarcinoma cases.

Materials and Methods

Case selection and tissue microarray construction. We obtained archived, formalin-fixed, paraffin-embedded tissue from surgically resected (with curative intent) lung cancer specimens (lobectomies and pneumonectomies) containing tumor and adjacent normal epithelium tissues from the Lung Cancer Specialized Program of Research Excellence Tissue Bank at The University of Texas M.D. Anderson Cancer Center, which has been approved by the Institutional Review Board. The tissue had been collected from 1997 to 2001, and the tissue specimens were histologically examined and classified using the 2004 WHO classification system (2). We selected 317 NSCLC tissue samples (201 adenocarcinomas and 116 squamous cell carcinomas) for our tissue microarrays. Tissue microarrays were constructed using triplicate 1-mm diameter cores per tumor, and each core included central, intermediate, and peripheral tumor tissue. Detailed clinical and pathologic information, including demographics, smoking history (never and ever smokers), and smoking status (never, former, and current), clinical and pathologic tumor-node-metastasis stage, overall survival duration, and time to recurrence, were available for most cases (Supplementary Table S1). Patients who had smoked at least 100 cigarettes in their lifetime were defined as smokers, and smokers who quit smoking at least 12 mo before their lung cancer diagnosis were defined as former smokers. Tumors were pathologic tumor-node-metastasis stages I to IV according to the revised International System for Staging Lung Cancer (23).

To assess the immunohistochemical expression of estrogen receptor α and β and progesterone receptor markers in the nonmalignant respiratory epithelium adjacent to lung tumors, we selected whole histology sections containing tumor and adjacent lung tissue from 64 adenocarcinomas that were included in our tissue microarrays.

Immunohistochemical staining and evaluation. The following antibodies against estrogen receptors α and β and progesterone receptor were purchased: (*a*) estrogen receptor α-1, clone 6F11, Novocastra, Leica Microsystems, Inc.; (*b*) estrogen receptor α-2, clone 6F11, Chemicon, Millipore Corporate; (*c*) estrogen receptor α-3, clone HC20, Santa Cruz Biotechnology, Inc.; (*d*) estrogen receptor α-4, clone 1D5, Lab Vision Corporation; (*e*) estrogen receptor β-1, clone H150, Santa Cruz Biotechnology; (*f*) estrogen receptor β-2, clone 14C8, GeneTex, Inc.; and (*g*) progesterone receptor, clone SP2, Lab Vision Corporation. Details on immunohistochemistry conditions and characteristics of the antibodies are listed in Supplementary Table S2. Immunohistochemical staining was done as follows: 5-μmol/L formalin-fixed, paraffin-embedded tissue sections were deparaffinized, hydrated, heated in a steamer for 10 min with 10 mmol/L sodium citrate (pH 6.0) for antigen retrieval, and washed in Tris buffer. Peroxide blocking was done with 3% H_2O_2 in methanol at room temperature for 15 min, followed by 10% fetal bovine serum in TBS-Tween for 30 min. The slides were incubated with primary antibody at an ambient temperature for 60 min for all antibodies; the exception was estrogen receptor β 14C8 (estrogen receptor β-2), which was incubated overnight at 4°C, washed with PBS, and incubated with biotin-labeled secondary antibody (Envision Dual Link+, DAKO) for 30 min. Staining for the slides was developed with 0.05% 3′,3-diaminobenzidine tetrahydrochloride, which had been freshly prepared in 0.05 mol/L Tris buffer (pH 7.6) containing 0.024% H_2O_2, and then, the slides were counterstained with hematoxylin, dehydrated, and mounted. Formalin-fixed, paraffin-embedded normal breast tissue was used as the positive control. For the negative control, we used the same specimens used for the positive controls but

replaced the primary antibody with PBS. For each antibody, we did titration experiments using a relatively wide range of antibody concentrations (1:50, 1:100, 1:200, and 1:500), including the concentration suggested by the manufacturer. The selection of the antibody dilution was based on the consistency in the expression in the breast cancer and normal tissues with known estrogen receptor and progesterone receptor expression used as control.

In addition, to study the correlation between the immunohistochemical expression of estrogen receptors and their ability to detected different isoforms of the proteins by Western blot, we examine by immunohistochemistry, formalin-fixed, paraffin-embedded cell lines pellets and by Western blot protein obtained from seven NSCLC cell lines (H157, H3255, HCC78, H1359, H1666, and H1174). For both techniques, we used the same antibodies that we used to examine the tissue specimens, except estrogen receptor α-2, which recently was replaced by the new manufacturer by a new clone (clone E115, Millipore; Supplementary Table S2). For Western blot analysis, cells were lysed with radioimmunoprecipitation assay buffer (50 mmol/L Tris-HCl, pH 7.4; 150 mmol/L NaCl; 1 mmol/L EDTA; 1% Triton X-100; 1% sodium deoxycholate; 0.1% SDS; 1 mmol/L phenylmethylsulfonylfluoride; 1 mmol/L Na_3VO_4; and protease inhibitor tablet; Roche Diagnostics), and 75 µg of protein was separated by SDS-PAGE. A new blot was run for each antibody to eliminate any potential cross-reactivity due to incomplete stripping of previous antibodies. Antibodies were diluted in 4% milk in Tris-buffered saline tween-20 (TBST) and incubated overnight at 4°C.

Two observers (M.G. Raso and I.I. Wistuba) jointly quantified the immunohistochemical expression of estrogen receptors and progesterone receptor using light microscopy (original magnification, ×20). Both nuclear and cytoplasmic expressions were quantified using a four-value intensity score (0, 1+, 2+, and 3+) and the percentage (0-100%) of reactivity. We defined the intensity categories as follows: 0, no appreciable staining; 1+, barely detectable staining in epithelial cells compared with the stromal cells; 2+, readily appreciable staining; and 3+, dark brown staining of cells. Next, an expression score was obtained by multiplying the intensity and reactivity extension values (range, 0-300).

EGFR mutation analysis. Exons 18 to 21 of *EGFR* were PCR-amplified using intron-based primers, as previously described (24, 25). Approximately 200 microdissected, formalin-fixed, paraffin-embedded cells were used for each PCR amplification. All PCR products were directly sequenced using the Applied Biosystems PRISM dye terminator cycle sequencing method. All sequence variants were confirmed by independent PCR amplifications from at least two independent microdissections and DNA extraction, and the variants were sequenced in both directions, as previously reported (24, 25).

Statistical analysis. The immunohistochemical expression and clinicopathologic data were summarized using standard descriptive statistics and frequency tabulations. BLiP plots were generated to summarize the distribution of estrogen receptor and progesterone receptor expressions. Associations between the marker expression and patients' clinical and demographic variables (including age, sex, smoking history, histology type, and pathologic stage) were assessed using appropriate methods, including the χ^2 or Fisher's exact test for categorical variables, and Wilcoxon rank sum or Kruskal-Wallis test for continuous variables. The Spearman rank r was used to estimate the correlation between immunohistochemistry markers. Kaplan-Meier survival curves for patient overall survival and recurrence-free survival were also generated. The log-rank test was used to identify the difference between the patient groups for overall and recurrence-free survival. For univariate and multivariate analyses for immunohistochemical expressions, the Cox proportional hazard model was used. Two-sided $Ps < 0.05$ were considered statistically significant.

Results

Correlation of expression by estrogen receptor antibodies. We examined four commercially available antibodies against estrogen receptor α: two using the same clone (6F11) and two antibodies against estrogen receptor β (Supplementary Table S2). Using the scores of expression generated from all NSCLCs, we analyzed the correlation of the expression in the malignant cells for the four estrogen receptor α and the two estrogen receptor β antibodies tested. All four of the estrogen receptor α antibodies showed nuclear staining, and two of the four antibodies also detected expression in the cytoplasm of malignant cells (estrogen receptor α-3, clone HC20, and estrogen receptor α-4, clone 1D5). Both estrogen receptor α clones 6F11 antibodies (estrogen receptors α-1 and α-2) against the full length of the protein, obtained from two different companies, showed only nuclear staining. Using the scores of expression, all four of the estrogen receptor α antibodies significantly correlated with each other at nuclear expression (Spearman rank correlation, $r = 0.32$-0.48; $P < 0.0001$; Supplementary Table S3). However, when only tumors expressing nuclear estrogen receptor α using any antibody were examined, no correlation in the expression using these four antibodies was observed (Supplementary Table S4). Similarly, although significant correlation was detected in the staining of the two estrogen receptor α antibodies (estrogen receptor α-3 and -4) showing cytoplasmic expression ($r = 0.43$; $P < 0.0001$) using the scores, no correlation was detected when only tumors showing cytoplasmic estrogen receptor α expression were examined. There was no statistically significant correlation between both of the estrogen receptor β antibodies examined in their nuclear expression; although they significantly correlated at their cytoplasmic expression, the r was very low ($r = 0.17$; $P = 0.005$).

To assess the ability of the antibodies to detect full length and isoforms of the estrogen receptor proteins, we examined their expression by immunohistochemistry (cell lines pellets) and Western blot in a panel of seven NSCLC cell lines. We identified that, by Western blot, all four estrogen receptor α and both estrogen receptor β antibodies used detect their full-length proteins of 66 and 60 kDa, respectively (data not shown). In addition, antibodies estrogen receptor α-1 and α-3 detect other isoforms of 46 kDa in NSCLC cell lines. We did not find correlation between the immunohistochemical nuclear and cytoplasmic expression and the patterns of isoforms detected by Western blot.

Frequency of estrogen receptor and progesterone receptor expression in NSCLC specimens by histology. We analyzed the frequency of any estrogen receptor and progesterone receptor immunohistochemical expression (positive cases, score > 0) for each antibody tested by NSCLC tumor histology, and the data are summarized in Table 1. Representative microphotographs of the expression of estrogen receptor and progesterone receptor with some of the antibodies tested are shown in Fig. 1. Estrogen receptors and progesterone receptor were detected in the nucleus of malignant cells by all of the corresponding antibodies tested. However, when expressed, the percentage of malignant cells showing staining was low in general, with an average percentage of positive expression of 19% (range, 2-90%) for estrogen receptor α-1 nuclear; 13% (range, 2-93%) for estrogen receptor α-2 nuclear; 21% (range, 1-60%) and 19% (range, 3-73%) for estrogen receptor α-3 nuclear and cytoplasmic, respectively; and 11% (range, 3-97%) and 7% (range, 3-30%) for estrogen receptor α-4 nuclear and cytoplasmic, respectively. The average percentages of positive cells expressing estrogen receptor β were 37% (range, 3-90%) and

Table 1. Frequency of estrogen receptor and progesterone receptor immunohistochemical expression in NSCLC tissue specimens

Marker	Location	ADCA		SCC		P
		No. of cases	Positive, n (%)	No. of cases	Positive, n (%)	
ERα-1	Nucleus	187	20 (11)	109	2 (2)	0.0048
	Cytoplasm	187	0	108	0	—*
ERα-2	Nucleus	186	84 (45)	110	23 (21)	<0.0001
	Cytoplasm	185	1 (<1)	111	0	1.000
ERα-3	Nucleus	191	16 (8)	114	0	0.0007
	Cytoplasm	190	92 (48)	114	37 (33)	0.0064
ERα-4	Nucleus	185	74 (40)	109	25 (23)	0.0028
	Cytoplasm	185	35 (19)	109	18 (17)	0.6043
ERβ-1	Nucleus	189	102 (54)	112	66 (59)	0.4022
	Cytoplasm	189	185 (98)	112	110 (98)	1.0000
ERβ-2	Nucleus	174	83 (48)	100	31 (31)	0.0069
	Cytoplasm	172	37 (22)	100	16 (16)	0.2685
PR	Nucleus	177	103 (58)	112	78 (70)	0.05
	Cytoplasm	176	0	112	0	—*

NOTE: Any expression score > 0 is considered positive.
Abbreviations: ADCA, adenocarcinomas; SCC, squamous cell carcinoma; ER, estrogen receptor; PR, progesterone receptor.
*Not tested.

37% (range, 3-97%) for estrogen receptor β-1 nuclear and cytoplasmic, respectively, and 13% (range, 1-77%) and 24% (range, 3-67%) for estrogen receptor β-2 nuclear and cytoplasmic, respectively.

Although there are important variations in the frequency of expression between the nuclear estrogen receptor α antibodies tested, adenocarcinoma histology showed significantly higher frequency of expression than squamous cell carcinomas for all estrogen receptor α antibodies (P < 0.0001-0.048; Table 1). For nuclear expression of estrogen receptor β, the data obtained with both antibodies tested were relatively consistent, and the adenocarcinoma histology showed a significantly higher frequency of expression than the squamous cell carcinoma did with the estrogen receptor β-2 antibody (P = 0.0069). Two of the estrogen receptor α (estrogen receptor α-3 and α-4) and both estrogen receptor β antibodies also detected estrogen receptor expression in the cytoplasm of NSCLC cells (Table 1). Although the estrogen receptor β-2 antibody detected protein expressed in the cytoplasm of a subset of NSCLCs, the estrogen receptor β-1 antibody detected expression in nearly all of the tumors. Cytoplasmic expression, only for the estrogen receptor α-3 antibody, was significantly higher in adenocarcinomas when compared with squamous cell carcinomas (P = 0.0064).

In the NSCLC tissues, progesterone receptor expression was frequently detected in the nuclei of malignant cells only. Squamous cell carcinoma histology showed a marginally significant higher frequency of expression than that of the adenocarcinomas (P = 0.05; Table 1).

Correlation between estrogen receptor and progesterone receptor expression in NSCLC and patients' clinicopathologic features. We correlated expression of estrogen receptors and progesterone receptor for each antibody tested with the patients' clinicopathologic characteristics, including histology, gender, tobacco history, and tumor-node-metastasis pathologic stage using the expression score as a continuous variable. Using this type of analysis, adenocarcinoma histology also showed a statistically significant higher nuclear expression for all estrogen receptor α antibodies and for the estrogen receptor β-2 antibody than squamous histology (Table 2). Of great interest was the fact that the NSCLC tissues obtained from females and never smokers showed statistically significant higher expression of nuclear estrogen receptor α and β for several of the antibodies used (Table 2). No correlations between the expression of progesterone receptor and the clinicopathologic characteristics were found.

We did overall survival and recurrence-free survival analyses to determine the expression of estrogen receptors and progesterone receptor for each antibody tested by using specimens from 317 patients with NSCLC with a median follow-up of 6.1 years for overall survival and 4.2 years for recurrence-free survival. No association was detected between the expression of estrogen receptor and progesterone receptor and overall survival. Of interest, any expression of cytoplasmic estrogen receptor α, using estrogen receptor α-4 antibody, and nuclear estrogen receptor β, using the estrogen receptor β-1 antibody, conferred to patients a significantly worse recurrence-free survival in the univariate and multivariate analysis (Fig. 2; Table 3). However, only the cytoplasmic expression of estrogen receptor α-4 correlated with worse recurrence-free survival when dichotomized score was being used (hazard ratio, 1.77; 95% confidence interval, 1.11-2.81; P = 0.0156; Table 3).

Correlation between estrogen receptor and progesterone receptor expression in NSCLC and tumor EGFR mutation status. Among 182 adenocarcinoma cases, EGFR mutations of the tyrosine kinase domain (exons 18-21) were detected in 31 (17%) cases. Most (88%) EGFR mutations were detected in the exons 19 and 21, and we did not find correlation between the location of the mutation and estrogen receptor α and β expression. We correlated the estrogen receptor and progesterone receptor scores and any expression (positive cases, score > 0) with EGFR mutation status. Interestingly, EGFR mutant adenocarcinomas showed statistically significant higher expression than wild-type tumors of nuclear estrogen receptor α, cytoplasmic estrogen receptor α, and nuclear estrogen receptor β when tested with
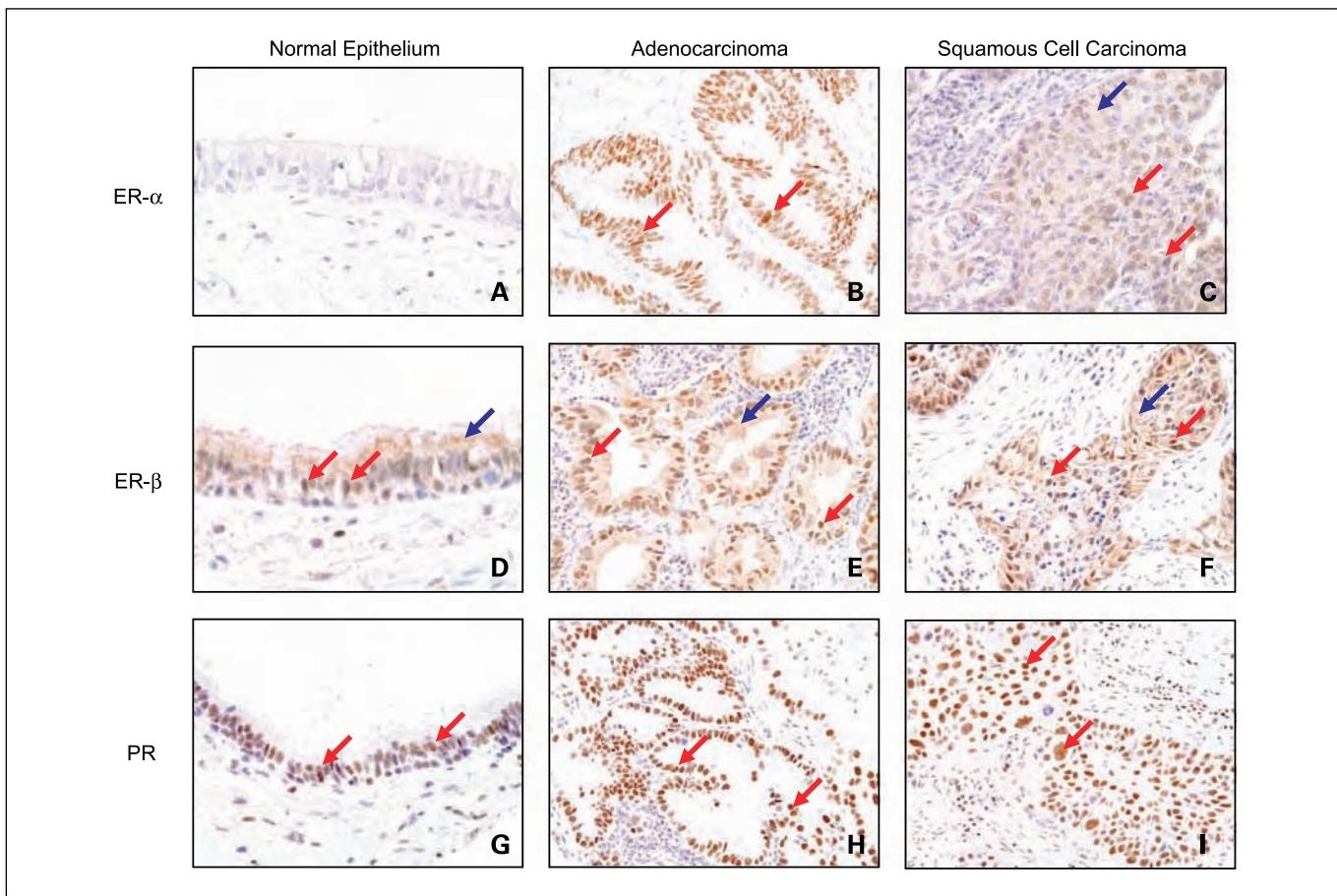
Fig. 1. Microphotographs showing representative examples of immunohistochemical expression of estrogen receptors α (*A-C*) and β (*D-F*) and progesterone receptor (*G-I*) in tissue specimens of NSCLC tumors and bronchial epithelium (*A, D* and *G*) adjacent to adenocarcinomas. The two NSCLC histologies are represented: adenocarcinoma (*B, E* and *H*), and squamous cell carcinoma (*C, F* and *I*). For adenocarcinoma, we show examples of high and low percentage of malignant cells expressing estrogen receptors and progesterone receptor. Estrogen receptor α and β expressions are shown using antibodies estrogen receptors α-4 and β-1, respectively. Red and blue arrows, examples of nuclear expression and cytoplasmic expressions, respectively. Normal epithelia show nuclear expression of estrogen receptor β and progesterone receptor and cytoplasmic expression of estrogen receptor β. Adenocarcinomas show nuclear expression for all three markers and cytoplasmic staining for estrogen receptor β. Squamous cell carcinomas show nuclear staining for progesterone receptor and nuclear and cytoplasmic for estrogen receptor β. Original magnification, ×200 for normal epithelium and ×400 for tumors.

antibodies estrogen receptor α-3, estrogen receptor α-4, and estrogen receptor β-1, respectively (Table 4; Fig. 3). Because there was a higher incidence of *EGFR* mutation in lung adenocarcinoma cases from patients with a history of never smoking, Asian ethnicity, or female characteristics (data not shown), we adjusted the effects of age, gender, smoking history, ethnicity, and pathologic stage in the correlation of estrogen receptor α and β with *EGFR* mutation status. After linear regression analysis,

Table 2. Significant correlations between immunohistochemical expression of estrogen receptor and progesterone receptor and NSCLC patients' clinicopathologic features

Estrogen receptor	Histology	Gender	Tobacco history
	ADCA (*n* = 201) > SCC (*n* = 116)	Female (*n* = 167) > Male (*n* = 150)	Never (*n* = 54) > Ever (*n* = 262)
ERα-1 nucleus	0.0048	0.0051	NS
ERα-2 nucleus	<0.0001	0.0109	0.0006
ERα-3 nucleus	0.0015	NS	0.0242
ERα-4 nucleus	0.0004	0.0148	0.0044
ERβ-1 nucleus	NS	NS	0.0290
ERβ-2 nucleus	0.0016	0.044	NS

NOTE: Estrogen receptor and progesterone receptor were tested using expression score.
Abbreviation: NS, not significant.

Table 3. Multivariate recurrence-free survival analysis using Cox regression model in NSCLC patients

Variable	HR	95% CI of HR		P
		Lower limit	Upper limit	
ER as continuous variable				
ERα-4 cytoplasm	1.05	1.01	1.08	0.0068
ERβ-1 nucleus	1.01	1.00	1.02	0.0034
Stage II vs I	1.90	1.14	3.18	0.0145
Stage III/IV vs I	3.17	1.98	5.08	<0.0001
ER dichotomized				
ERα-4 cytoplasm: >0 vs 0	1.77	1.11	2.81	0.0156
ERβ-1 nucleus: >0 vs 0	1.36	0.91	2.05	0.1388
Stage II vs I	1.79	1.08	2.99	0.0250
Stage III/IV vs I	3.13	1.97	4.99	<0.0001

NOTE: With only significant covariates.
Abbreviations: 95% CI, 95% confidence interval; HR, hazard ratio.

all the significant correlations remained statistically significant. There was no correlation between progesterone receptor expression and *EGFR* mutation status.

Estrogen receptor and progesterone receptor immunohistochemical expression in the lung respiratory airway adjacent to adenocarcinoma cases. To characterize the pattern of expression of estrogen receptor and progesterone receptor in the respiratory airway field in patients with lung cancer, we selected 64 adenocarcinoma cases (35 females and 29 males; 19 never smoked, 13 current smokers, and 32 former smokers), and we studied the immunohistochemical expression of estrogen receptor α and β and progesterone receptor in the respiratory cells lining the small bronchi (*n* = 35 cases), bronchioles (*n* = 83 cases), and alveoli exhibiting type II cells hyperplastic changes (*n* = 15 cases) using the same semiquantitative scoring system used in the tissue microarrays. For estrogen receptors, we tested the estrogen receptor α-4 and estrogen receptor β-1 antibodies. From each case, we used immunohistochemistry to examine whole tissue sections from a mean of three different paraffin blocks (range, 3-6) containing tumor and adjacent normal lung tissue. We found that estrogen receptor α was not expressed in the airway epithelium adjacent to lung adenocarcinomas, including epithelial samples from 21 positive tumors (Supplementary Table S5). In contrast, estrogen receptor β was widely expressed in the cytoplasm of respiratory cells: 91% of bronchi, 84% of bronchioles, and 29% of the hyperplastic alveoli. Estrogen receptor β nuclear immunostaining was found less frequently: 5% of bronchi, 10% of bronchioles, and none of the hyperplastic alveolar cells. Noticeably, in the bronchial cells, we identified two patterns of cytoplasmic immunostaining: a homogeneous staining in all types of bronchial cells and heterogeneous staining comprising only ciliated cells with mainly supranuclear or apical expression (Fig. 1). Progesterone receptor was found in the nucleus of 56% of bronchi, 61% of bronchioles, and 33% of hyperplastic alveoli. Of interest, there was a high level of correlation (28 of 33 comparisons, 85%) between the expression of progesterone receptor in the normal epithelium and the corresponding tumors. Twenty (95%) of 21 cases with progesterone receptor positive in the normal epithe-

lium were detected in patients with tumors that also expressed this receptor.

Discussion

Estrogen receptors α and β frequently expressed in our NSCLC cases, and estrogen receptor α expression distinguished a subset of NSCLC that has defined clinicopathologic and genetic features. Although the immunohistochemical expression of estrogen receptors α and β has been reported in tumor tissue specimens from surgically resected NSCLCs, the data on the fraction of tumors expressing estrogen receptor are still controversial. Previous studies on estrogen receptor α immunohistochemical expression in formalin-fixed and paraffin-embedded NSCLC specimens using six different antibodies identified nuclear expression in malignant cells in frequencies that ranged from none (10, 14) to 18% (26) and 38% (12). Similarly, in other studies, the frequency of estrogen receptor α cytoplasmic expression in NSCLC ranged from 0% to 3% (12, 26) to 35% (11) and 73% (27). In the current study, using four different commercially available estrogen receptor α antibodies, we also identified a wide range of percentages in the frequency of NSCLCs exhibiting any expression of estrogen receptor α in the nucleus (7-54%) and in the cytoplasm (0-42%) of tumor cells. However, in our study, when the scores of immunohistochemical expression were analyzed as continuous variables, all of the estrogen receptor α antibodies significantly correlated with each other at nuclear and cytoplasmic locations.
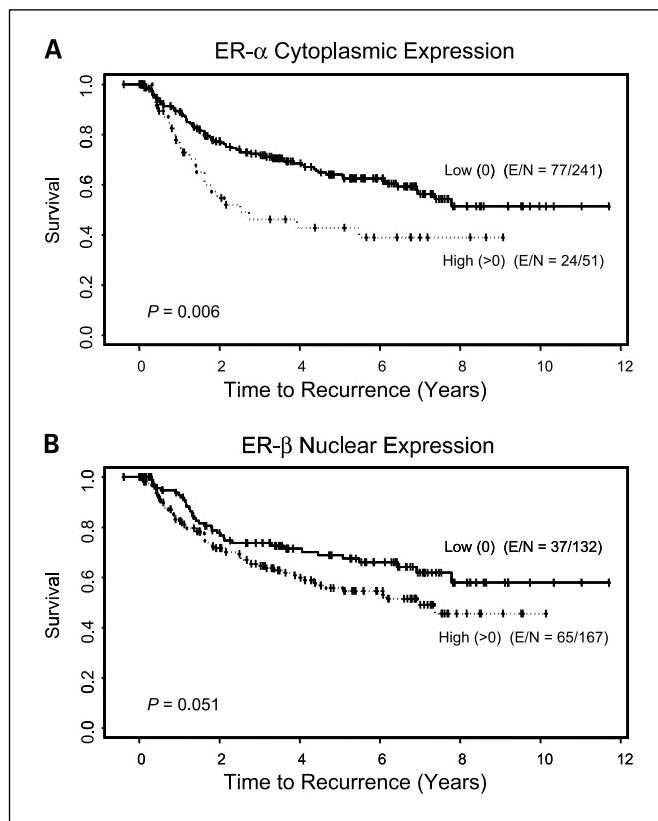
Fig. 2. Kaplan-Meier curves showing recurrence-free survival of NSCLC patients for estrogen receptor α cytoplasmic (*A*) and estrogen receptor β nuclear (*B*) expression.

Table 4. Significant correlations between immunohistochemical expression of estrogen receptor and *EGFR* mutation status in adenocarcinoma

ER expression by antibody	*EGFR* mutation status		*P*
	Wild-type, *n* positive/total (%)	Mutant, *n* positive/total (%)	
ERα-3 nucleus	9/146 (6)	7/28 (25)	0.0016
ERα-3 cytoplasm	68/146 (47)	21/27 (78)	0.0029
ERα-4 nucleus	50/143 (35)	18/27 (67)	0.0020
ERα-4 cytoplasm	20/143 (14)	13/27 (48)	<0.0001
ERβ-1 nucleus	70/145 (48)	22/27 (82)	0.00015

NOTE: Estrogen receptor and progesterone receptor were tested using expression score.

A similar situation is observed when the estrogen receptor β immunohistochemical expression data are examined in NSCLC. Several previous studies, using six different antibodies, have reported frequencies of estrogen receptor β expression in tumors with a wide range of percentages at the nuclear location, 0% (9), 34% to 47% (10, 12, 14), and 61% to 84% (9, 11), but not in the cytoplasm of malignant cells, wherein most of the studies have shown no reactivity (9, 10, 12, 14); some expression was seen in a small number of cases (6) or low frequency of expression in a large number of cases (10%; ref. 11). In the present study, using two antibodies, any estrogen receptor β nuclear expression was detected in about half (56% and 42%) of the NSCLCs, and cytoplasmic expression was found in a wider range (20-98%) of our cases. We do not have a definitive explanation to the high levels of expression of estrogen receptor β in NSCLC cells in our study and the discordance with previous reports. However, immunohistochemical analysis has shown the distri-

bution of estrogen receptor β to be much more widespread than estrogen receptor α (28–30). Several studies have reported that estrogen receptor β immunohistochemical expression is frequently detected in the nucleus and cytoplasm of normal respiratory cells (28). Although expression has been questioned by suggestions that this observation is based on nonspecific binding produced by unpurified antibodies (31), multiple reports have shown the presence of a nonnuclear pool of estrogen receptors in normal and malignant cells (32–35). Yang et al. (35) used one of the same estrogen receptor β antibodies that we used (estrogen receptor β-1) and showed mitochondrial localization of this receptor in several normal human and murine cells, suggesting a role for estrogen receptor β receptor in the cytoplasm of cells. Our finding of high frequency of estrogen receptor β expression, using estrogen receptor β-1 antibody, in the cytoplasm of normal respiratory cells from our lung adenocarcinoma patients are consistent with these findings.
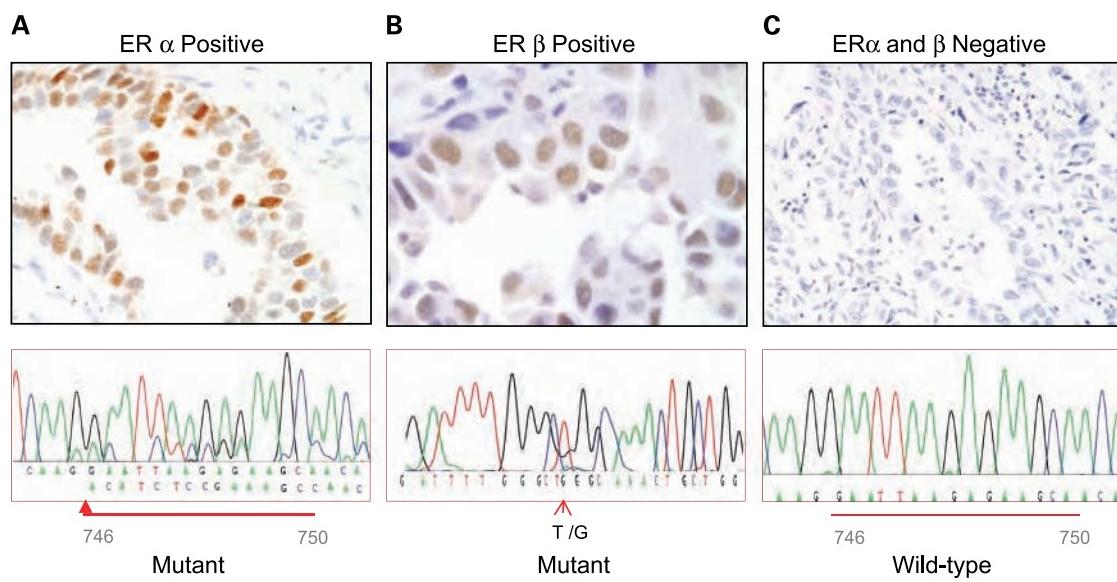
Fig. 3. Representative examples of estrogen receptor α and β immunohistochemical expression (*top figures*) and EGFR mutations (*bottom figures*) in lung adenocarcinomas. *A*, estrogen receptor α (antibody estrogen receptor α-4) positive in the nucleus of malignant cells and sequencing chromatograms showing the presence of mutant form of *EGFR* (15-bp deletion in exon 19). Arrow, in-frame deletion mutation sequence. *B*, estrogen receptor β (antibody estrogen receptor β-1) positive in the nucleus of malignant cells and sequencing chromatograms showing the presence of mutant *EGFR* (L858R point mutation in exon 21). Arrow, CTG to CGG mutation. *C*, estrogen receptors α and β (same antibodies than as *A* and *B*) with negative expression in the malignant cells and sequencing chromatograms showing the presence of wild-type form of *EGFR* exon 19. Line, sequences 746 to 750.

Several discrepancies were observed when we compared our results with those published previously (6, 9, 10, 12) using the same antibodies, especially for estrogen receptor α. For example, our estrogen receptor α-3 antibody, raised against the COOH-terminus region of the protein, detected nuclear and cytoplasmic expressions in 54% and 42% of our NSCLC cases, respectively. Using this antibody, nuclear expression was reported in a small number of NSCLC tumors by Stabile et al. (6) and in none of the 130 tumors examined by Kawai et al. (10). At the cytoplasmic location of malignant cells, both studies reported positive immunostaining (6, 10), with up to 73% of cases in the study done by Kawai et al. (10).

Why these inconsistent results on the immunohistochemical expression of estrogen receptor α and β occur raises a very important question. Clearly, the reasons for the inconsistent results include the use of different antibodies manufactured from different clones and by different companies. Indeed, some of these antibodies have been made against different parts of the protein: full length, NH_2-terminus, and COOH-terminus regions. It has been shown that several mRNA splicing variants of estrogen receptor α have been detected in lung cancer cell lines, and antibodies raised against epitopes in the deleted exons of estrogen receptor may give conflicting results (6). Although a number of estrogen receptor α mRNA variants have been reported, in most cases, wild-type mRNA estrogen receptor α is coexpressed along with splicing variants (36). In our study, we identified a high level of discordance in the frequency and location in the malignant cells of the expression of estrogen receptor α and estrogen receptor β antibodies examined. By examining NSCLC cell lines using immunohistochemistry and Western blot, we did not find correlation between the patterns of immunostaining (nuclear versus cytoplasmic) and the expression protein isoforms. In addition, it is important to note that there are multiple criteria reported to assess estrogen receptor α and β positivity in NSCLC tissues. Although most studies considered different expression intensity levels (usually a scale 0-3+) at nuclear and cytoplasmic locations combined with the percentage of malignant cells expressing a given intensity, the cutoff levels of expression vary significantly between studies (e.g., 1+ in >10% of cells; 1+ in 1-25% of cells; >50% of cells; score 0-8, etc.; refs. 6, 9–12, 14, 26).

Because there were different levels of estrogen receptor α and β immunohistochemical expression detected using different antibodies in ours and the previous studies (6, 9–12, 14, 26), we correlated the expression of estrogen receptor using all of the antibodies we tested with the patients' clinicopathologic features and the tumors' *EGFR* mutation status. The evaluation of multiple antibodies for estrogen receptor expression adds strength to our findings. In our study, we analyzed the immunohistochemical scores as continuous and dichotomized variables, and a significantly higher expression of nuclear estrogen receptor α was detected with all four antibodies tested in adenocarcinoma than squamous cell carcinoma histology, three of the four antibodies tested in tumors obtained from females compared with males and from people who had never smoked compared with smokers. The biological implication of the large range of positive lung cancer cells for estrogen receptor and progesterone receptor is unknown. The cutoff that we selected (positive, >0) to study correlation with clinicopathologic features was based in the fact that a large number of tumors showed low percentage of malignant cells expressing these re-

ceptors. The two previous studies reporting estrogen receptor α nuclear expression in NSCLC, which examined a relatively large series of cases, did not address differences of expression based on histology types or patients' clinicopathologic features (12, 26). In the NSCLC tissues that we reviewed, higher expression of estrogen receptor β correlated significantly with tumor adenocarcinoma histology and the patients' female sex for estrogen receptor β-1 antibody and correlated with the patients' history of never smoking with the estrogen receptor β-2 antibody.

Few studies have shown inconsistent results on whether estrogen receptor expression is biased to any sex using different types of specimens and assays (6–9). Schwartz et al. (9), using a different antibody than ours, reported that NSCLCs obtained from females were 46% less likely to have estrogen receptor β–positive tumors than males in a multivariate analysis. In addition, mRNA expression of estrogen receptor α has been reported to be significantly higher in lung tumors from women than from men (8). In a small number of NSCLC tumor tissue specimens, estrogen receptor α and β gene transcripts have been found to be expressed in similar levels when comparing samples obtained from females and males (7). Adenocarcinoma of the lung, which shows a weaker association with tobacco smoking than with other types of lung cancer, is also found predominantly in women, suggesting a possible role for female hormones in the pathogenesis of this type of lung cancer (5).

In previous studies, estrogen receptor β expression in NSCLC tumors has been associated with improved survival (9–11), whereas the immunohistochemical expression of estrogen receptor α has been shown to be a poor prognostic factor (9). Thus, both estrogen receptors have been proposed to play opposite roles in cell proliferation, with estrogen receptor α promoting proliferation and estrogen receptor β having an antiproliferative effect (37, 38). In our study, we did not find a correlation between overall survival and recurrence-free survival and estrogen receptor β expression, but we did find that only the expression of cytoplasmic estrogen receptor α (using one antibody) conferred to patients a significantly worse recurrence-free survival, but not overall survival, in multivariate analysis.

Several studies have shown that estrogen signaling plays a role in the development of the epithelium in the lung and that estrogen could potentially promote lung cancer (6, 7, 13, 16, 17). In addition, antiestrogen drugs have been suggested to have a role in the therapy of lung cancer (6, 19). NSCLC cell lines and *in vivo* tumor xenografts have been shown to respond to estrogens, and tumor growth can be inhibited up to 40% by the antiestrogen fulvestrant (6). In the past few years, significant advances have been made in the development of new molecularly targeted agents for lung cancer (39). The identification of the subset of patients with NSCLC who will benefit with targeted therapy is a key element in the development of personalized treatment approaches in this disease. A pilot study on combined therapy using fulvestrant and gefitinib in advanced NSCLC has shown to be well tolerated and has shown some tumor responses (40). Our study results strongly suggest that NSCLC tumors obtained from patients with adenocarcinoma histology, female gender, and history of never smoking have a higher chance of expressing estrogen receptors and have the potential to respond positively to antiestrogen therapy.

While our manuscript was under review, Nose et al. (41), using one antibody against estrogen receptor α (same antibody as our estrogen receptor α-3) and estrogen receptor β (same antibody as our estrogen receptor β-1), reported in 447 surgically resected lung adenocarcinomas (stages I-IV) from Japanese patients that estrogen receptor β nuclear expression was significantly (*P* < 0.001) higher in tumors with *EGFR* mutation. They also reported that estrogen receptor β nuclear expression correlated with increasing disease-free survival (hazard ratio, 2.18; 95% confidence interval, 1.18-4.06; *P* = 0.014) in the patients with *EGFR* mutant tumors. Although, in this study, a different scoring system was used to asses estrogen receptor immunohistochemical expression and the expression of estrogen receptors α and β in the malignant cells was reported at cytoplamic and nuclear levels only, respectively, the finding of significant correlation of estrogen receptor β nuclear expression with *EGFR* mutation in lung adenocarcinoma agrees with our findings. Thus, our study is the first to report an association between *EGFR* mutation and estrogen receptor α expression in lung adenocarcinomas. Importantly, we have shown that the correlation between estrogen receptor expression and *EGFR* mutation is independent of the clinicopathologic features associated with both abnormalities, such as adenocarcinoma histology, female gender, and history of never smoking (42). Based on the interactions between estrogen receptor and EGFR-signaling pathways, there is evidence showing that targeting both pathways by using antiestrogens (fulvestrant) and EGFR tyrosine kinase inhibitors (gefitinib), the antitumor effect in *in vitro* and *in vivo* lung cancer models of the drug combination is higher than in treatment with each drug alone (19). Thus, our findings of an association between the activation of both pathways further strengthens the concept of combined antiestrogen and EGFR inhibitor therapy for a selected group of patients with lung adenocarcinoma.

Although progesterone receptor expression has been reported to be present in NSCLC cell lines and tumor specimens, the data are controversial like those for estrogen receptors (11, 12, 15, 43, 44). Of four studies reporting on immunohistochemical expression of progesterone receptor in surgical resected and formalin-fixed NSCLC tissue specimens using different antibodies, there were two studies that reported a relatively high frequency of progesterone receptor expression in tumors (39% and 47%; refs. 12, 15); the remaining two reports showed no expression (11, 44). In the present study, progesterone receptor was frequently (63%) detected in the nuclei of malignant NSCLC, with a trend to higher expression in squamous cell carcinoma histology. We did not find a correlation between progesterone receptor and any of the clinicopathologic charac-

teristics we studied, including survival. In contrast, Ishibashi et al. (12) reported that progesterone receptor immunohistochemical expression was higher in NSCLCs obtained from females and correlated with better overall survival in stages I to III tumors. In breast cancer, transcription of the progesterone receptor gene is well known to be regulated by estrogenic actions through estrogen receptors, and a positive progesterone receptor status is generally regarded as one of the markers of functional estrogenic pathways. In our study, we found no statistical correlation between progesterone receptor and any of the estrogen receptor antibodies studied. *In vitro* and *in vivo* studies have shown that administration of progesterone inhibits the growth of progesterone receptor-positive NSCLC cell lines, which is similar to what has been shown to happen in breast and endometrial carcinomas (12).

Lung cancer is believed to develop from a series of preneoplastic lesions in the respiratory mucosa, and these abnormalities are frequently extensive and multifocal throughout the respiratory epithelium, indicating a field-effect or field-cancerization phenomenon (45). Our findings of relatively frequent expression of nuclear progesterone receptor and lack of expression of estrogen receptor α in the normal epithelium adjacent to adenocarcinomas expressing these receptors suggest that progesterone receptor, but not estrogen receptor α expression, may represent a field-effect phenomenon. Of interest, all but one case with normal epithelium expression of progesterone receptor showed expression of this receptor in the corresponding tumor. The frequent finding of cytoplasmic estrogen receptor β in normal epithelium may represent a constitutive expression in normal respiratory cells and is probably not related to the carcinogenesis process (35).

In summary, our findings show that estrogen receptors α and β and progesterone receptor are frequently expressed in NSCLC, and estrogen receptor expression distinguishes a subset of NSCLC that has defined clinicopathologic and genetic features. In our study, there is a bias toward early stages of lung cancer because we examined surgically resected tumors, being most (98%) of them stages I to III. The frequency of estrogen receptor and progesterone receptor expression in advanced metastatic tumors need to be further examined. The correlation between estrogen receptor and *EGFR* mutation in lung adenocarcinoma suggests that it might be important to target both pathways simultaneously for lung cancer chemoprevention and therapy.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005; 55:74–108.

2. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Tumours of the lung. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, editors. Pathology and genetics: tumours of the lung, pleura, thymus and heart. Lyon: IARC; 2004, p. 9–124.

3. Brennan P, Bray I. Recent trends and future directions for lung cancer mortality in Europe. Br J Cancer 2002;87:43–8.

4. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. CA Cancer J Clin 2008;58:71–96.

5. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers—a different disease. Nat Rev Cancer 2007;7:778–90.

6. Stabile LP, Davis AL, Gubish CT, et al. Human non-small cell lung tumors and cells derived from normal lung express both estrogen receptor α and β and show biological responses to estrogen. Cancer Res 2002;62:2141–50.

7. Mollerup S, Jorgensen K, Berge G, Haugen A. Expression of estrogen receptors α and β in hu-

man lung tissue and cell lines. Lung Cancer 2002; 37:153–9.

8. Fasco MJ, Hurteau GJ, Spivack SD. Gender-dependent expression of α and β estrogen receptors in human nontumor and tumor lung tissue. Mol Cell Endocrinol 2002;188:125–40.

9. Schwartz AG, Prysak GM, Murphy V, et al. Nuclear estrogen receptor β in lung cancer: expression and survival differences by sex. Clin Cancer Res 2005;11:7280–7.

10. Kawai H, Ishii A, Washiya K, et al. Estrogen receptor α and β are prognostic factors in

non-small cell lung cancer. Clin Cancer Res 2005; 11:5084–9.

11. Skov BG, Fischer BM, Pappot H. Oestrogen receptor β over expression in males with non-small cell lung cancer is associated with better survival. Lung Cancer 2008;59:88–94.

12. Ishibashi H, Suzuki T, Suzuki S, et al. Progesterone receptor in non-small cell lung cancer—a potent prognostic factor and possible target for endocrine therapy. Cancer Res 2005;65:6450–8.

13. Marquez-Garban DC, Chen HW, Fishbein MC, Goodglick L, Pietras RJ. Estrogen receptor signaling pathways in human non-small cell lung cancer. Steroids 2007;72:135–43.

14. Wu CT, Chang YL, Shih JY, Lee YC. The significance of estrogen receptor β in 301 surgically treated non-small cell lung cancers. J Thorac Cardiovasc Surg 2005;130:979–86.

15. Su JM, Hsu HK, Chang H, et al. Expression of estrogen and progesterone receptors in non-small-cell lung cancer: immunohistochemical study. Anticancer Res 1996;16:3803–6.

16. Pietras RJ, Marquez DC, Chen HW, Tsai E, Weinberg O, Fishbein M. Estrogen and growth factor receptor interactions in human breast and non-small cell lung cancer cells. Steroids 2005;70:372–81.

17. Hershberger PA, Vasquez AC, Kanterewicz B, Land S, Siegfried JM, Nichols M. Regulation of endogenous gene expression in human non-small cell lung cancer cells by estrogen receptor ligands. Cancer Res 2005;65:1598–605.

18. Stabile LP, Siegfried JM. Estrogen receptor pathways in lung cancer. Curr Oncol Rep 2004; 6:259–67.

19. Stabile LP, Lyker JS, Gubish CT, Zhang W, Grandis JR, Siegfried JM. Combined targeting of the estrogen receptor and the epidermal growth factor receptor in non-small cell lung cancer shows enhanced antiproliferative effects. Cancer Res 2005;65:1459–70.

20. Sequist LV, Lynch TJ. EGFR tyrosine kinase inhibitors in lung cancer: an evolving story. Annu Rev Med 2008;59:429–42.

21. Herbst RS, Bunn PA, Jr. Targeting the epidermal growth factor receptor in non-small cell lung cancer. Clin Cancer Res 2003;9:5813–24.

22. Fujimoto N, Wislez M, Zhang J, et al. High expression of ErbB family members and their ligands in lung adenocarcinomas that are sensitive to inhibition of epidermal growth factor receptor. Cancer Res 2005;65:11478–85.

23. Mountain CF. Revisions in the international system for staging lung cancer. Chest 1997; 111:1710–7.

24. Shigematsu H, Lin L, Takahashi T, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 2005;97: 339–46.

25. Tang X, Shigematsu H, Bekele BN, et al. EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. Cancer Res 2005;65: 7568–72.

26. Lau SK, Chu PG, Weiss LM. Immunohistochemical expression of estrogen receptor in pulmonary adenocarcinoma. Appl Immunohistochem Mol Morphol 2006;14:83–7.

27. Kawai H, Ishii A, Washiya K, et al. Combined overexpression of EGFR and estrogen receptor α correlates with a poor outcome in lung cancer. Anticancer Res 2005;25:4693–8.

28. Taylor AH, Al-Azzawi F. Immunolocalisation of oestrogen receptor β in human tissues. J Mol Endocrinol 2000;24:145–55.

29. Jarvinen TA, Pelto-Huikko M, Holli K, Isola J. Estrogen receptor β is coexpressed with ERα and PR and associated with nodal status, grade, and proliferation rate in breast cancer. Am J Pathol 2000;156:29–35.

30. Skliris GP, Carder PJ, Lansdown MR, Speirs V. Immunohistochemical detection of ERβ in breast cancer: towards more detailed receptor profiling? Br J Cancer 2001;84:1095–8.

31. Saunders PT. Does estrogen receptor β play a significant role in human reproduction? Trends Endocrinol Metab 2005;16:222–7.

32. Collins P, Webb C. Estrogen hits the surface. Nat Med 1999;5:1130–1.

33. Razandi M, Oh P, Pedram A, Schnitzer J, Levin ER. ERs associate with and regulate the production of caveolin: implications for signaling and cellular actions. Mol Endocrinol 2002;16:100–15.

34. Speirs V, Carder PJ, Lane S, Dodwell D, Lansdown MR, Hanby AM. Oestrogen receptor β: what it means for patients with breast cancer. Lancet Oncol 2004;5:174–81.

35. Yang SH, Liu R, Perez EJ, et al. Mitochondrial localization of estrogen receptor β. Proc Natl Acad Sci U S A 2004;101:4130–5.

36. Herynk MH, Fuqua SA. Estrogen receptor mutations in human disease. Endocr Rev 2004;25: 869–98.

37. Lindberg MK, Moverare S, Skrtic S, et al. Estrogen receptor (ER)-β reduces ERα-regulated gene transcription, supporting a "ying yang" relationship between ERα and ERβ in mice. Mol Endocrinol 2003;17:203–8.

38. Stettner M, Kaulfuss S, Burfeind P, et al. The relevance of estrogen receptor-β expression to the antiproliferative effects observed with histone deacetylase inhibitors and phytoestrogens in prostate cancer treatment. Mol Cancer Ther 2007;6:2626–33.

39. Lynch TJ, Bonomi PD, Butts C, et al. Novel agents in the treatment of lung cancer: Fourth Cambridge Conference. Clin Cancer Res 2007; 13:s4583–8.

40. Traynor AM, Schiller JH, Stabile LP, et al. Pilot study of gefitinib and fulvestrant in the treatment of post-menopausal women with advanced non-small cell lung cancer. Lung Cancer 2008.

41. Nose N, Sugio K, Oyama T, et al. Association between estrogen receptor-β expression and epidermal growth factor receptor mutation in the postoperative prognosis of adenocarcinoma of the lung. J Clin Oncol 2009;27:411–7.

42. Shigematsu H, Gazdar AF. Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers. Int J Cancer 2006;118: 257–62.

43. Kaiser U, Hofmann J, Schilli M, et al. Steroid-hormone receptors in cell lines and tumor biopsies of human lung cancer. Int J Cancer 1996;67: 357–64.

44. Di Nunno L, Larsson LG, Rinehart JJ, Beissner RS. Estrogen and progesterone receptors in non-small cell lung cancer in 248 consecutive patients who underwent surgical resection. Arch Pathol Lab Med 2000;124:1467–70.

45. Wistuba II. Genetics of preneoplasia: lessons from lung cancer. Curr Mol Med 2007;7:3–14.

HER Family Receptor Abnormalities in Lung Cancer Brain Metastases and Corresponding Primary Tumors

Menghong Sun,[1] Carmen Behrens,[2] Lei Feng,[3] Natalie Ozburn,[1] Ximing Tang,[2] Guosheng Yin,[3] Ritsuko Komaki,[4] Marileila Varella-Garcia,[5] Waun Ki Hong,[2] Kenneth D. Aldape,[1] and Ignacio I. Wistuba[1,2]

Abstract

Purpose: To compare the characteristics of deregulation of HER receptors and their ligands between primary tumor and corresponding brain metastases of non–small cell lung carcinoma (NSCLC).

Experimental Design: Fifty-five NSCLC primary tumors and corresponding brain metastases specimens were examined for the immunohistochemical expression of epidermal growth factor receptor (EGFR), phosphorylated EGFR, Her2, Her3, and phosphorylated Her3, and their ligands EGF, transforming growth factor-α, amphiregulin, epiregulin, betacellulin, heparin-binding EGFR-like growth factor, neuregulin (NRG) 1, and NRG2. Analysis of *EGFR* copy number using fluorescence *in situ* hybridization and mutation by PCR-based sequencing was also done.

Results: Metastases showed significantly higher immunohistochemical expression of EGF (membrane: brain metastases 66.0 versus primary tumors 48.5; $P = 0.027$; nucleus: brain metastases 92.2 versus primary tumors 67.4; $P = 0.008$), amphiregulin (nucleus: brain metastases 53.7 versus primary tumors 33.7; $P = 0.019$), phosphorylated EGFR (membrane: brain metastases 161.5 versus primary tumors 76.0; $P < 0.0001$; cytoplasm: brain metastases 101.5 versus primary tumors 55.9; $P = 0.014$), and phosphorylated Her3 (membrane: brain metastases 25.0 versus primary tumors 3.7; $P = 0.001$) than primary tumors did. Primary tumors showed significantly higher expression of cytoplasmic transforming growth factor-α (primary tumors 149.8 versus brain metastases 111.3; $P = 0.008$) and NRG1 (primary tumors 158.5 versus brain metastases 122.8; $P = 0.006$). In adenocarcinomas, a similar high frequency of *EGFR* copy number gain (high polysomy and amplification) was detected in primary (65%) and brain metastasis (63%) sites. However, adenocarcinoma metastases (30%) showed higher frequency of *EGFR* amplification than corresponding primary tumors (10%). Patients whose primary tumors showed *EGFR* amplification tended to develop brain metastases at an earlier time point.

Conclusions: Our findings suggest that NSCLC brain metastases have some significant differences in HER family receptor–related abnormalities from primary lung tumors.

Lung cancer is the leading cause of cancer-related deaths in the United States (1). Lung cancer includes several histologic types, the most frequently occurring of which (~80%) are two types of non–small cell lung carcinoma (NSCLC): adenocarcinoma and squamous cell carcinoma (1). The brain is one of the main sites of metastasis in patients with lung cancer: brain metastasis has an incidence of up to 60% in patients with lung adenocarcinoma (2–5). The median survival for lung cancer patients with brain metastasis is usually 3 to 6 months (5, 6). The use of systemic chemotherapy and cranial irradiation is unsuccessful in the treatment of NSCLC brain metastasis (2, 7), and this in turn has motivated the search for new therapeutic strategies for this disease.

During the past few years, significant advances have been made in the development of new molecularly targeted agents for lung cancer (8). One example of such targets is the epidermal growth factor receptor (EGFR) that belongs to the HER family tyrosine kinase (TK) receptors composed of four homologous cell membrane receptors, including Her2 and Her3 (9). These three receptors are activated by nine known ligands, including EGF, transforming growth factor-α (TGF-α), amphiregulin, epiregulin, betacellulin, heparin-binding EGFR-like growth factor (HB-EGF), neuregulin (NRG) 1, and NRG2 (10–12). Deregulation of HER receptors, especially EGFR, seems to play an important role in the pathogenesis and

---

Translational Relevance

Brain metastasis occurs in up to 60% of non–small cell lung carcinomas, and there is little information on the molecular differences between primary tumor and metastases. Our findings indicate that non–small cell lung carcinoma brain metastases have some significant differences in HER family receptor–related abnormalities from primary lung tumors. These differences could be related to tumor progression and may cause diverse responses to epidermal growth factor receptor and other HER receptor–targeted therapy of primary and metastatic tumor sites.

---

progression of NSCLC (13). In lung cancer cells, the constitutive activation of EGFR is achieved by several mechanisms, including increased production of ligands, increased levels of the receptor, and mutation of *EGFR* (13–15).

Small-molecule inhibitors that target the TK domain of the *EGFR* produce tumor responses in ∼10% of patients with advanced NSCLC that has progressed despite prior chemotherapy (13, 16). However, the brain is still a frequent site of disease recurrence in NSCLC patients after an initial response to treatment with EGFR TK inhibitors (TKI), regardless of disease control in the lungs (3). Activating mutations in the *EGFR* TK domain, an increased *EGFR* copy number, and increased EGFR protein expression have been associated with a favorable response to treatment with EGFR TKIs (13, 16). Previous reports showed that metastatic NSCLC brain tumors respond to EGFR TKIs (17, 18). However, it is still unclear whether the abnormalities of EGFR and its related receptors differ in metastases compared with primary NSCLC tumors.

The identification of differences in the deregulation of HER ligands and receptors in NSCLC primary tumors and corresponding metastases may explain differences in the response to HER receptor–targeted therapy in this tumor type, and these differences need to be considered in the analysis of predictive biomarker used in clinical trials. In the present study, we investigated the immunohistochemical expression of three HER receptors (EGFR, Her2, and Her3) implicated in the pathogenesis of NSCLC cancer and eight of their nine known ligands (EGF, TGF-α, amphiregulin, epiregulin, betacellulin, HB-EGF, NRG1, and NRG2) in 55 paired primary and brain metastasis NSCLCs. In addition, we compared *EGFR* copy number and mutation abnormalities using fluorescence *in situ* hybridization (FISH) and PCR-based sequencing analyses, respectively.

Materials and Methods

NSCLC tissue specimens. We obtained archived formalin-fixed paraffin-embedded material from surgically resected specimens from 55 NSCLC patients with primary lung cancers and corresponding brain metastases containing tumor tissues and collected between 1988 and 2002. These cases were selected based on the availability of enough archival tissue for the immunohistochemistry and FISH analyses. All specimens were from the lung cancer tissue bank at The University of Texas M. D. Anderson Cancer Center, which is approved by the M. D. Anderson institutional review board. After histologic examination,

tissue microarrays were constructed using three 1-mm-diameter cores per tumor.

Detailed clinical and pathologic information, including demographic, pathologic tumor-node-metastasis staging, overall survival, and time of brain metastasis occurrence, was obtained for all patients (Table 1). Pathologic tumor-node-metastasis stage had been determined for lung cancers according to the revised International System for Staging Lung Cancer (19) at time of primary tumor surgery with curative intent. In all cases, the NSCLC brain metastases were solitary, and 11 patients also developed metastases at other brain sites over a median period of 12 mo (range, <1-27 mo). Forty-four (80%) of 55 patients developed clinically detectable brain metastases after primary lung cancer surgical resection (metachronous tumors; median, 13 mo; range, <1-94 mo); in 11 (20%) patients, the brain metastases were detected at the same time as the lung tumors (synchronous tumors), and they were surgically removed before (median, <1 mo range, <1-11 mo) the primary lung cancer surgery.

Immunohistochemical staining and evaluation. For our analysis, antibodies against the following proteins were purchased and used: EGF (dilution, 1:50; EMD Biosciences), amphiregulin (dilution, 1:150; Lab Vision), TGF-α (dilution, 1:150; EMD Biosciences), epiregulin (dilution, 1:10; R&D Systems), betacellulin (dilution, 1:10; R&D Systems), HB-EGF (dilution, 1:10; R&D Systems), NRG1 (dilution, 1:10; R&D Systems), NRG2 (dilution, 1:50; Abcam, Inc.), EGFR (clone 31G7; dilution, 1:100; Zymed), p-EGFR Tyr[1086] (dilution, 1:100; Invitrogen), Her2 (dilution, 1:100; Dako), Her3 (dilution, 1:50; GenTex), and phosphorylated Her3 (p-Her3; dilution, 1:100; Cell Signaling). Immunohistochemical staining was done using 5-μm-thick tissue microarray histologic sections as previously described (20). The immunohistochemical protein expression was quantified, using white light microscopy with ×20 magnification, by two experienced thoracic pathologists (M.S. and I.I.W.) blinded to clinical and other molecular variables. All markers were examined for membrane, cytoplasm, and nucleus localization in tumor cells. As previously described (21–23), immunohistochemical expression was quantified using a three-value intensity score (0, 1+, 2+, and 3+) for all markers, except for membrane EGFR and phosphorylated EGFR (p-EGFR), for which a four-value intensity score (0, 1+, 2+, 3+, and 4+) and the percentage (0-100%) of the extent of reactivity were used. Next, expression scores

---

Table 1. Summary of clinicopathologic features of 55 NSCLC patients with primary tumors and corresponding brain metastases

Characteristic	Number	%
Tumor histology		
Adenocarcinoma	40	73
Squamous cell carcinoma	13	23
Large cell carcinoma	1	2
Adenosquamous carcinoma	1	2
Age (y)		
≤60	30	55
>60	25	45
Gender		
Female	19	35
Male	36	65
Pathologic stage*		
I	13	24
II	10	18
III	18	33
IV [†]	14	25

* The staging was done at time of surgical resection of the primary lung tumor.
[†] In 11 cases, the brain metastases were surgically removed before the primary tumor.
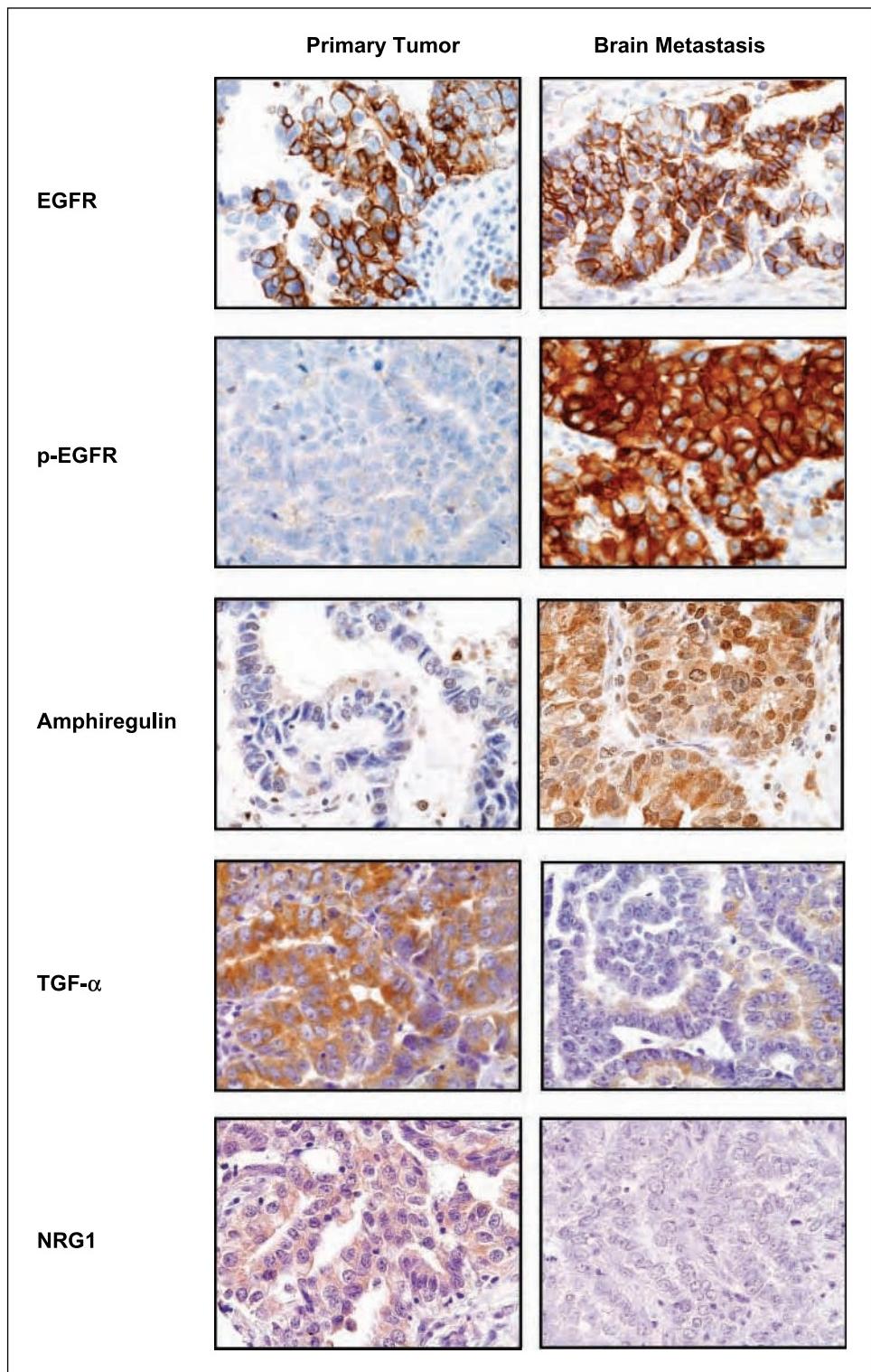
Fig. 1. Representative microphotographs of immunohistochemical expression of EGFR and p-EGFR and the ligands amphiregulin, TGF-α, and NRG1 in primary tumors and corresponding brain metastases. Magnification, ×400. All markers showed protein expression (*brown staining*) in tumor cells from primary and/or metastasis sites at the membrane and cytoplasm levels. Amphiregulin showed also nuclear expression in malignant cells.

were obtained by multiplying the intensity and reactivity extension values (range, 0-300 for all markers, except for membrane EGFR and p-EGFR with a range of 0-400). For each case of primary tumor and metastasis, the immunohistochemical expression of the markers was averaged using the cores available per tumor site.

EGFR *FISH analysis*. We analyzed the gene copy number per cell using the LSI *EGFR* SpectrumOrange/CEP 7 SpectrumGreen Probe

(Abbott Molecular), as previously described (24, 25). Tumor specimens were classified into six FISH strata according to the frequency of cells with each *EGFR* gene copy number and referred to the chromosome 7 centromere, as follows: (*a*) disomy (3 or 4 copies in <10% of cells), (*b*) low trisomy (3 copies in 10% to <40% of cells and 4 copies in <10% of cells), (*c*) high trisomy (3 copies in ≥40% of cells and 4 copies in <10% of cells), (*d*) low polysomy (4 copies in 10% to <40% of cells),
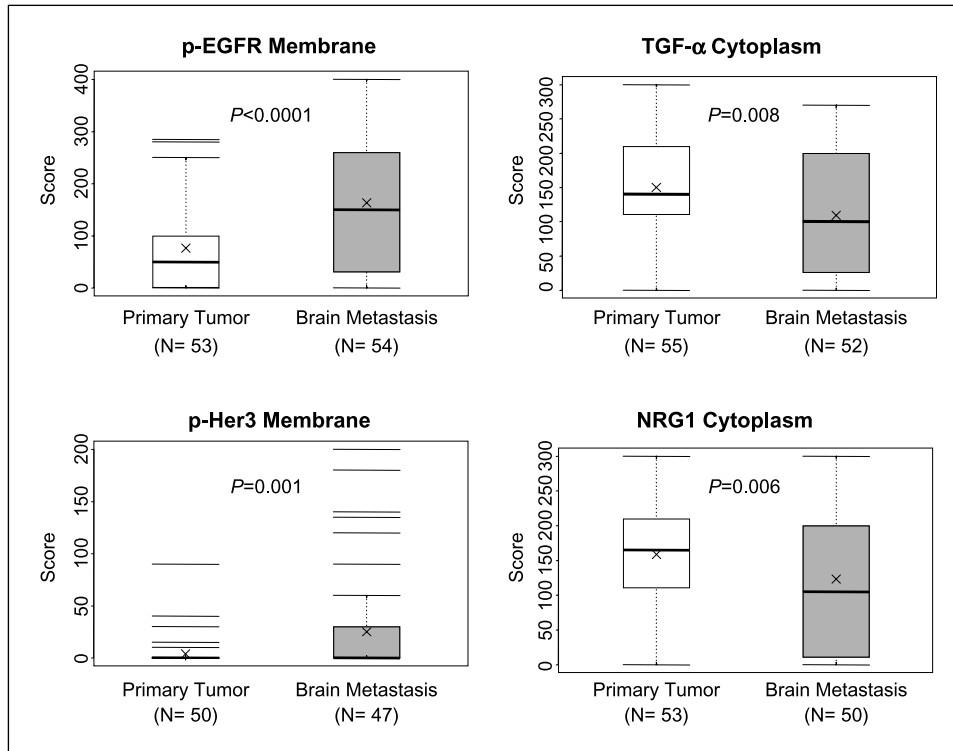
Fig. 2. Box plots showing scores of immunohistochemical expression of p-EGFR, TGF-α, p-Her3, and NRG1 markers comparing primary tumors with corresponding brain metastases. *P* values comparing primary tumors and brain metastasis are shown for all comparisons. Bars, median; x, mean score.

(*e*) high polysomy (≥ 4 copies in $\geq 40\%$ of cells), and (*f*) gene amplification (presence of loose or tight *EGFR* gene clusters with ≥ 4 copies, *EGFR* gene to CEP 7 ratio ≥ 2, or 15 copies of *EGFR* per cell in $\geq 10\%$ of cells). The high polysomy and gene amplification categories were considered to indicate high *EGFR* copy number (*EGFR* FISH positive), and the other categories were considered to indicate no significant increase in the *EGFR* copy number (*EGFR* FISH negative), as previously described (24, 25). For each case of primary tumor and metastasis, the FISH *EGFR* copy number was quantified by counting cells representing each core available per tumor site.

EGFR mutation analysis. Exons 18 to 21 of *EGFR* were PCR amplified using intron-based primers as previously described (26). Approximately 200 microdissected formalin-fixed paraffin-embedded cells were used for each PCR amplification. All PCR products were directly sequenced using the Applied Biosystems Prism dye terminator cycle sequencing method. All sequence variants were confirmed by independent PCR amplifications from at least two independent microdissections and DNA extraction, and the variants were sequenced in both directions, as previously reported (26).

Statistical analysis. Data were summarized using standard descriptive statistics and frequency tabulations. Associations between the marker expression and patients' clinical demographic variables, including age, sex, histology type, and pathologic stage, were assessed using appropriate methods, including the χ^2 test or Fisher's exact test for categorical variables and the Wilcoxon rank sum test or the Kruskal-Wallis test for continuous variables. The Wilcoxon signed rank test was used to test the differences in biomarker expression between primary lung tumors and brain metastases. Cox proportional hazard models were used for univariate analysis of time to metastasis according to biomarker expression. Hazard ratios (HR) with 95% confidence intervals (95% CI) and *P* values are reported. All tests were two sided. *P* values <0.05 were considered to indicate statistical significance.

Results

Immunohistochemical expression of HER receptors and ligands in NSCLC primary tumors and corresponding brain metastases.

Most markers, including the ligands EGF, amphiregulin, TGF-α, epiregulin, betacellulin, NRG1, and NRG2 and the receptors EGFR, p-EGFR, Her2, Her3, and p-Her3, showed protein expression in tumor cells from primary and metastasis sites at the membrane and cytoplasm levels (Fig. 1; Supplementary Table S1). Of those, EGF, amphiregulin, epiregulin, NRG1, NRG2, p-EGFR, and p-Her3 showed also nuclear expression in malignant cells (Fig. 1; Supplementary Fig. S1). The ligand HB-EGFR expressed only in the cytoplasm of cancer cells. Although showing overlapping, brain metastases had significantly higher immunohistochemical expression scores of EGF (membrane: metastasis 66.0 versus primary 48.5; *P* = 0.027; nucleus: metastasis 92.2 versus primary 67.4; *P* = 0.008), amphiregulin (nucleus: metastasis 55.4 versus primary 33.7; *P* = 0.019), p-EGFR (membrane: metastasis 161.5 versus primary 76.0; *P* < 0.0001; cytoplasm: metastasis 101.5 versus primary 55.9; *P* = 0.014), and p-Her3 (membrane: metastasis 25.0 versus primary 3.7; *P* = 0.001) than did corresponding primary tumors (Fig. 2; Supplementary Table S1). Only the protein expression score of TGF-α (primary 149.8 versus metastasis 111.3; *P* = 0.008) and NRG1 (primary 158.5 versus metastasis 122.8; *P* = 0.006) at the cytoplasmic level was significantly higher in malignant cells from primary tumors than in brain metastasis cells (Fig. 2; Supplementary Table S1).

EGFR copy number analysis by FISH in NSCLC primary tumors and corresponding brain metastases. Overall, the presence of high frequency of gain in *EGFR* copy number (FISH positive: high polysomy and amplification; Fig. 3) was similar in NSCLC primary (34 of 55, 62%) and brain metastasis (35 of 55, 64%) sites (Table 2). Although a relatively lower frequency of high polysomy was detected in metastases than in primary tumors (33% versus 47%), brain metastases showed a nonsignificant higher frequency of *EGFR* amplification than

corresponding primary tumors did (31% versus 15%; $P = 0.53$). In adenocarcinomas ($n = 40$ cases), a similar frequency of gain in *EGFR* copy number was detected in primary tumors (65%) and corresponding metastases (63%). However, brain metastases of lung adenocarcinoma showed a nonsignificant higher frequency of *EGFR* amplification than primary lung tumors (30% versus 10%; $P = 0.53$). Although a higher frequency in *EGFR* copy number gain was detected in brain metastases (62%) than primary tumors (46%) among squamous cell carcinomas (Table 2), the data were difficult to interpret because of the small number of cases available for analysis. The concordance of *EGFR* copy number abnormalities between both tumor sites was higher for the cases with primary and

metastasis tumors clinically detected as synchronous lesions (11 of 11, 100%) than those diagnosed as metachronous tumors (34 of 44, 77%).

A relatively high level of concordance (46 of 55, 84%) for gain in *EGFR* copy number gain was found between primary tumors and metastases (Supplementary Table S2). Sixteen (29%) paired primary/metastasis cases were *EGFR* FISH negative in both sites, whereas 30 (55%) paired cases showed gain in *EGFR* copy number at both tumor sites. Discordance in *EGFR* copy number status was detected in nine cases (16%); in six of these, brain metastasis sites had a gain in copy number, whereas primary tumors were FISH negative. The levels of concordance for high polysomy (15 of 30, 50%) and
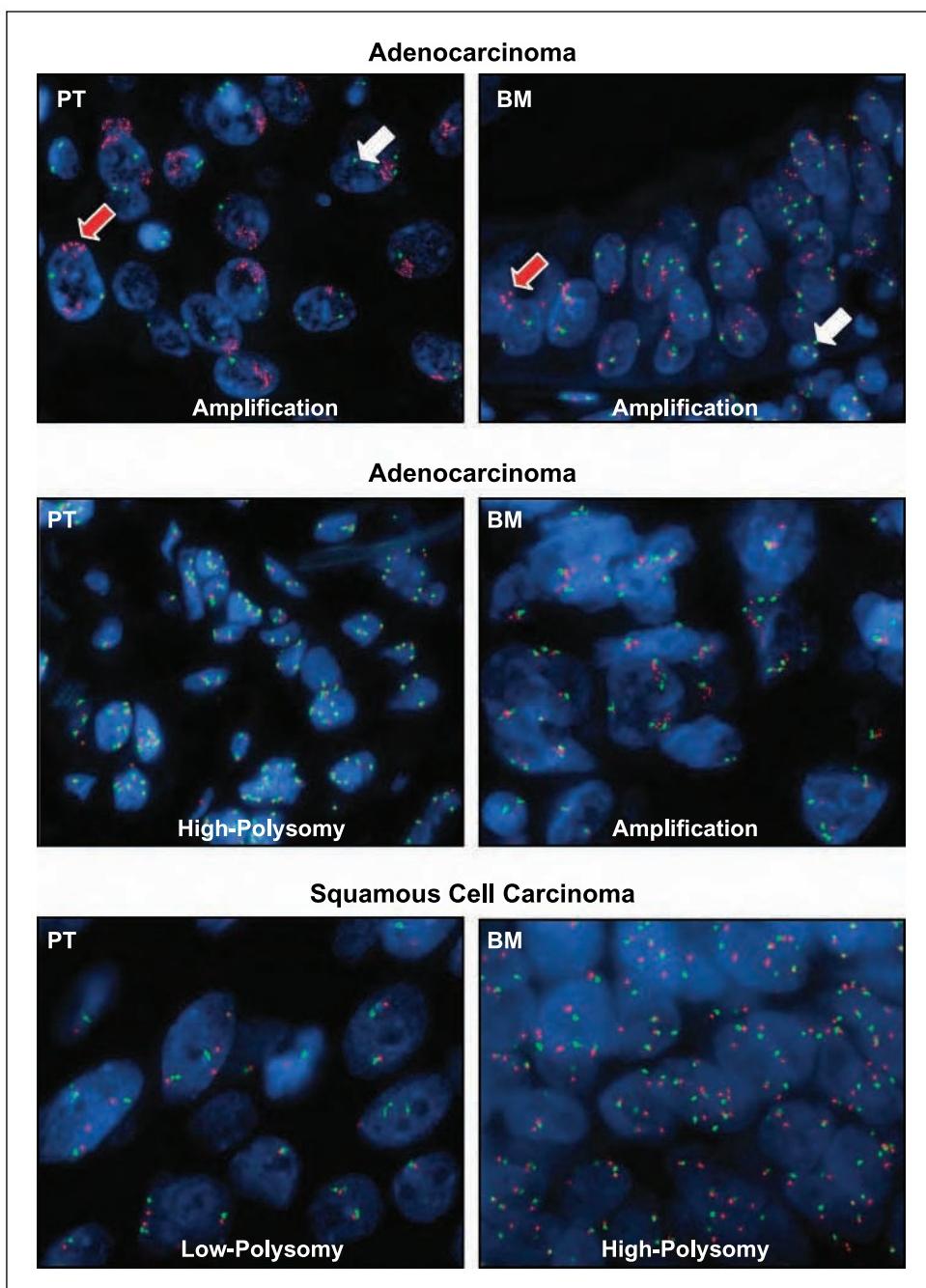
Fig. 3. Representative microphotographs of FISH showing *EGFR* copy number in primary tumors (*PT*) and corresponding brain metastases (*BM*). Magnification, ×1,000. Red signals (*red arrows*) represent *EGFR* gene copies and green signals (*white arrows*) represent the chromosome 7 centromere probe. Cell nuclei stained blue with 4′,6-diamidino-2-phenylindole. High polysomy is defined by ≥4 copies in ≥40% of cells, and gene amplification by the presence of loose or tight *EGFR* gene clusters and a ratio of *EGFR* gene to chromosome of 2 or 15 copies of *EGFR* per cell in 10% of the analyzed cells.

Table 2. *EGFR* copy number by FISH in 55 NSCLC primary and corresponding brain metastases by tumor histology

Copy number categories	Adenocarcinoma (*n* = 40)		Squamous cell carcinoma (*n* = 13)		Total (*N* = 55)*	
	Primary	Metastasis	Primary	Metastasis	Primary	Metastasis
	n (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)	*n* (%)
FISH negative	14 (35)	15 (37)	7 (54)	5 (38)	21 (40)	20 (36)
Disomy	1 (3)	2 (5)	0 (0)	0 (0)	1 (2)	2 (4)
Trisomy	2 (5)	2 (5)	1 (8)	0 (0)	3 (6)	2 (4)
Low polysomy	11 (28)	11 (27)	6 (46)	5 (39)	17 (31)	16 (29)
FISH positive	26 (65)	25 (63)	6 (46)	8 (62)	34 (62)	35 (64)
High polysomy	22 (55)	13 (33)	2 (15)	5 (39)	26 (47)	18 (33)
Amplification	4 (10)	12 (30)	4 (31)	3 (23)	8 (15)	17 (31)

*One adenosquamous carcinoma and one large cell carcinoma showed *EGFR* high polysomy in the primary tumors and amplification in the brain metastasis specimen.

amplification (6 of 18, 33%) were low when primary tumors and corresponding brain metastases were compared. To assess heterogeneity in *EGFR* copy number abnormalities in the different cores examined per tumor sample, we examined in paired primary tumors and metastases from 14 cases the level of concordance between the level of the most advanced gene copy number abnormality detected in all cores examined per tumor site. In primary tumors, the concordance was 100% (22 comparisons), and in the brain metastases was 90% (27 of 30 comparisons).

Two consecutive brain metastasis samples were available for analysis from each of five adenocarcinoma cases. The length of time between the consecutive brain metastasis was 1, 6, 10, 16, and 31 months. In all comparisons, paired consecutive brain metastasis specimens showed identical *EGFR* copy number status. One pair was FISH negative and the other four were FISH positive. FISH-positive specimens included two pairs showing *EGFR* high polysomy and two showing gene amplification.

We correlated *EGFR* copy number status with the immunohistochemical expression of the markers at both tumor sites. The only associations detected were that *EGFR* FISH-positive primary tumors and brain metastases showed significantly higher protein expression scores of nuclear p-EGFR ($P = 0.018$) and cytoplasmic Her2 ($P = 0.015$) than FISH-negative tumors. In contrast, FISH-positive tumors showed lower expression of betacellulin ($P = 0.0293$) than FISH-negative tumors.

EGFR *mutations in NSCLC primary tumors and corresponding brain metastases.* We successfully amplified and sequenced DNA obtained from primary tumors and metastases samples from 42 cases, including 30 adenocarcinomas, for exons 19 and 21 of *EGFR*, which harbor >90% of TK activating mutations of the gene (27). We detected only one case with *EGFR* mutation (exon 19, point mutation TTA2239-2240CCA, Leu747Pro), which was present in both sites, primary and metastasis, in adenocarcinoma obtained from a patient female and never smoker.

Correlation between immunohistochemical expression of markers and EGFR *copy number and time to brain metastasis.* We investigated the correlation between the immunohistochemical expression of the markers examined and *EGFR* copy number abnormalities in primary lung tumors and the time to brain metastasis development. In this analysis, we included only the 44 patients whose brain metastases were

diagnosed after surgical resection of the primary tumor. Overall, the median time to brain metastasis for all 44 patients was 1.23 years (95% CI, 0.89-1.62 years). The median time to brain metastasis development for patients with adenocarcinoma was 1.43 years (95% CI, 0.96-2.04 years) and that for patients with squamous cell carcinoma was 0.89 years (95% CI, 0.63 to not available). Using the Cox proportional hazard regression models, we identified that adenocarcinoma, compared with squamous cell carcinoma, was significantly correlated with a longer time to brain metastasis occurrence ($P = 0.009$; HR, 0.347; 95% CI, 0.157-0.769), whereas EGFR membrane protein expression scores ($P = 0.025$; HR, 1.003; 95% CI, 1.000-1.006) and *EGFR* amplification (versus no amplification) ($P = 0.0039$; HR, 3.492; 95% CI, 1.494-8.162) were significantly correlated with a shorter time to brain metastasis development. None of these markers was shown in the multivariate analysis to be statistically significant predictors of metastasis development. However, in the multivariate analysis, when we analyzed time to brain metastasis development in the subset of 39 patients who developed a single brain metastasis (excluding the 5 patients who developed two consecutive metastases in the brain) after primary tumor diagnosis, we found that adenocarcinoma, compared with squamous cell carcinoma, was significantly correlated with a longer time to brain metastasis occurrence ($P = 0.031$; HR, 0.373; 95% CI, 0.152-0.917). In contrast, *EGFR* amplification (versus no amplification) was significantly correlated ($P = 0.0033$; HR, 4.452; 95% CI, 1.645-12.053) with a shorter time to brain metastasis development.

Discussion

In NSCLC, overexpression and activation of EGFR, Her2, and Her3 are well-known phenomena (9, 13). However, to the best of our knowledge, the overexpression of those TK receptors has not been previously reported in NSCLC brain metastasis. In this study, we have described for the first time higher levels of immunohistochemical expression of EGFR, p-EGFR, Her2, Her3, and p-Her3 in a series of NSCLC brain metastases using tissue microarray specimens. Interestingly, we found that the expression of phosphorylated forms of EGFR and Her3 proteins at the cytoplasmic and membrane level of malignant cells was significantly increased in brain metastasis compared with expression in corresponding primary lung tumors. Although

these data need to be validated in a larger set of specimens, these findings are consistent with the notion that activation of the EGFR and Her3 pathways is important in the progression and metastasis of lung cancer (13, 16). Similarly to brain metastasis, we recently showed that in *EGFR* mutant lung adenocarcinomas, p-EGFR immunohistochemical expression was significantly increased in nine lymph node metastases compared with expression in corresponding primary tumors (20).

It is known that the receptors of the HER family are activated after binding to ligands or peptide growth factors (10, 12). Ligand binding induces clustering of HER family receptors and produces subsequent autophosphorylation of cytoplasmic tyrosine residues (10, 12). There are 11 HER ligands identified, and they can be divided into four groups based on the receptor binding specificity: (*a*) exclusive EGFR binding: EGF, amphiregulin, TGF-α, and epigen; (*b*) EGFR and Her4 binding: betacellulin, HB-EGFR, and epiregulin; (*c*) Her3 and Her4 binding: NRG1 and NRG2; and (*d*) exclusive Her4 binding: NRG3 and NRG4 (10, 12). No ligand binding Her2 has been identified (10, 12). In our study, we examined the protein expression of eight of nine ligands that bind to EGFR and Her3 receptors (10, 12). Of these, TGF-α (28, 29), epiregulin (30), and amphiregulin (31) are frequently expressed in primary NSCLC tumors, and EGF, betacellulin, HB-EGF, and NRG1 have been shown to be expressed in NSCLC cell lines (32–34). However, none of them has been characterized in primary lung tumors and corresponding brain metastasis. In addition, to the best of our knowledge, there is no report of the expression of NRG2 in lung cancer. We found that, compared with the corresponding primary tumors, NSCLC brain metastases had significantly higher immunohistochemical expression of membrane and nuclear EGF and of nuclear amphiregulin—ligands associated with activation of EGFR dimmers (10). These findings are consistent with the concomitant high level of overexpression of p-EGFR in the NSCLC brain metastasis that we studied and indicate the presence of an autocrine secretion mechanism of these ligands. In contrast to EGF and amphiregulin, the cytoplasmic expression of TGF-α and NRG1, which bind to EGFR and Her3 receptors, respectively (10, 12), was significantly higher in malignant cells from primary tumors than in cells from brain metastases. Overexpression of TGF-α has been associated with the metastatic potential of NSCLC (32) and colon cancer (35) cell lines in favoring modifications of the tumor microenvironment conducive to metastasis, such as increasing angiogenesis.

In our study, we have identified that six of ligands, EGF, amphiregulin, epiregulin, NRG1, and NRG2, and two receptors, p-EGFR and p-Her3, had nuclear expression in malignant NSCLC cells. There is evidence that TK receptors, as well as their ligands, translocate into the nucleus via receptor-mediated endocytosis for degradation or to be recycled back to the cell surface (36–40). However, it now seems clear that these complexes reach into the cell nucleus where they participate directly in the control of cell proliferation, cell differentiation, and cell survival (40).

The current concept of metastasis development states that metastases are the result of tumor cells interacting with a specific organ microenvironment, also called the "seed and soil" hypothesis (41). Thus, the microenvironments of different organs, including the brain, are biologically unique and can explain the expression of HER receptors and ligands in the brain metastasis tissue specimens differing from expression in

the corresponding primary lung tumors. In addition, these observations have important implications for the development of molecularly targeted therapy in lung cancer patients. The fact that potential therapeutic targets (EGF, amphiregulin, TGF-α, NRG1, EGFR, and Her3) are expressed differently in metastases from corresponding primary tumors suggests that different molecular properties among tumor sites may influence differing responses to treatment and affect the levels of biomarkers that may be predictive of the response to treatment. Although immunohistochemical testing of EGFR has been shown not to be an optimal method for identifying patients who may respond to treatment with anti-EGFR drugs (16), there are preliminary data suggesting that the expression in tumor tissue of Her3 (15), amphiregulin (42), and TGF-α (31) correlates with sensitivity and resistance to EGFR TKI therapy. The immunohistochemical overexpression of Her3 in NSCLC tissue specimens has been correlated with EGFR TKI sensitivity (15). In contrast, increased expression of amphiregulin and TGF-α has been correlated with resistance to such therapy (31). In breast cancer, the transmembrane expression of neuroregulin has been correlated with improved survival in patients treated with Her2 inhibitor (43).

An increase in *EGFR* gene copy number, including high polysomy and gene amplification (as shown by FISH), has been detected in 22% of patients with surgically resected (stages I-IIIA) NSCLC (21). Higher frequencies (40-50%) of *EGFR* high copy number have been reported in patients with more-advanced metastatic NSCLC (stage IV; refs. 44–48). In the present study, we have identified even a higher frequency (62%) of gain in *EGFR* copy number in surgically resected primary NSCLC specimens from patients who developed brain metastases. Recently, we reported that a gain in *EGFR* gene copy number was detected in 74% of primary NSCLC tumors from patients who developed brain metastasis (25). Altogether, these data suggest a stepwise increase in the frequency of gain in *EGFR* copy number in primary tumors with increasing tumor stage and, more important, with the development of brain metastasis. Interestingly, in our cases, the presence of *EGFR* amplification, along with membrane EGFR protein overexpression, was significantly correlated with shorter time to brain metastasis development in the univariate analysis, further suggesting the important role of this genetic abnormality in the progression and metastasis of NSCLC.

Recently, we (20) and others (49) have shown that *EGFR* copy number gain, and specifically gene amplification, is a late phenomenon in the development of lung adenocarcinoma, appearing at invasive tumor stages and progressing in lymph node metastases, and that it is preceded by gene mutation. In the present study, we have expanded some of these observations to NSCLC brain metastasis. Although it was not statistically significant, we found that brain metastases of lung adenocarcinomas had a higher frequency of *EGFR* amplification than the corresponding primary tumors (30% versus 10%). Although a relatively high level (84%) of concordance for gain in *EGFR* copy number (when high polysomy and amplification were analyzed together) was detected when primary tumors and metastases were compared, there were nine discordant cases (16%), including six brain metastases that had increased copy numbers, whereas primary tumors did not. In contrast, we found that *EGFR* gene amplification had a low level of concordance (33%) when primary and metastatic

tumors were compared, indicating a high level of heterogeneity for this phenomenon. The distinct rate of *EGFR* gene amplification between primary tumors and corresponding brain metastases may support the influence of this phenomenon on differing responses to treatment and may affect the assessment of this specific biomarker for anti-EGFR therapy.

The low frequency of *EGFR* mutations in exons 19 and 21 detected in our series of 42 primary NSCLC and corresponding metastases examined, including 30 adenocarcinomas, did not allow us to compare differences between both tumor sites. The single case having *EGFR* mutation (exon 19, point mutation) in the primary tumor showed identical mutation in the metastasis. The concordance on *EGFR* mutation between primary tumors and brain metastases has been previously reported in NSCLC in Japanese patients (50).

In summary, our findings indicate that NSCLC brain metastases exhibit important differences in abnormalities related to the HER family receptors from primary lung tumors. These differences may cause different responses to EGFR and other HER receptor–targeted therapy of primary and metastatic tumor sites and suggest that the site of origin (primary versus metastasis) of the tumor specimen should be factored into the biomarker analyses in the clinical trials testing the efficacy of HER receptor inhibitor in patients with metastatic NSCLC. Although our series of cases is relatively small and restricted to one metastatic site per patient, the data strongly suggest that the analysis of both primary and metastasis tumor sites may be critical for the identification of novel therapeutic targets and corresponding predictive biomarkers in lung cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Tumours of the lung. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, editors. Pathology and genetics: tumours of the lung, pleura, thymus and heart. Lyon: IARC; 2004. p. 9–124.

2. Stuschke M, Eberhardt W, Pottgen C, et al. Prophylactic cranial irradiation in locally advanced non-small-cell lung cancer after multimodality treatment: long-term follow-up and investigations of late neuropsychologic effects. J Clin Oncol 1999;17:2700–9.

3. Omuro AM, Kris MG, Miller VA, et al. High incidence of disease recurrence in the brain and leptomeninges in patients with nonsmall cell lung carcinoma after response to gefitinib. Cancer 2005;103:2344–8.

4. Mamon HJ, Yeap BY, Janne PA, et al. High risk of brain metastases in surgically staged IIIA non-small-cell lung cancer patients treated with surgery, chemotherapy, and radiation. J Clin Oncol 2005;23:1530–7.

5. Chen AM, Jahan TM, Jablons DM, Garcia J, Larson DA. Risk of cerebral metastases and neurological death after pathological complete response to neoadjuvant therapy for locally advanced nonsmall-cell lung cancer: clinical implications for the subsequent management of the brain. Cancer 2007;109:1668–75.

6. Zabel A, Debus J. Treatment of brain metastases from non-small-cell lung cancer (NSCLC): radiotherapy. Lung Cancer 2004;45 Suppl 2:S247–52.

7. Lester JF, MacBeth FR, Coles B. Prophylactic cranial irradiation for preventing brain metastases in patients undergoing radical treatment for non-small-cell lung cancer: a Cochrane Review. Int J Radiat Oncol Biol Phys 2005;63:690–4.

8. Lynch TJ, Bonomi PD, Butts C, et al. Novel agents in the treatment of lung cancer: Fourth Cambridge Conference. Clin Cancer Res 2007;13:s4583–8.

9. Herbst RS. Review of epidermal growth factor receptor biology. Int J Radiat Oncol Biol Phys 2004; 59:21–6.

10. De Luca A, Carotenuto A, Rachiglio A, et al. The role of the EGFR signaling in tumor microenvironment. J Cell Physiol 2008;214:559–67.

11. Montero JC, Rodriguez-Barrueco R, Ocana A, Diaz-Rodriguez E, Esparis-Ogando A, Pandiella A. Neuregulins and cancer. Clin Cancer Res 2008;14: 3237–41.

12. Huang Z, Brdlik C, Jin P, Shepard HM. A pan-HER approach for cancer therapy: background, current status and future development. Expert Opin Biol Ther 2009;9:97–110.

13. Sequist LV, Lynch TJ. EGFR tyrosine kinase inhibitors in lung cancer: an evolving story. Annu Rev Med 2008;59:429–42.

14. Herbst RS, Bunn PA, Jr. Targeting the epidermal growth factor receptor in non-small cell lung cancer. Clin Cancer Res 2003;9:5813–24.

15. Fujimoto N, Wislez M, Zhang J, et al. High expression of ErbB family members and their ligands in lung adenocarcinomas that are sensitive to inhibition of epidermal growth factor receptor. Cancer Res 2005; 65:11478–85.

16. Ciardiello F, Tortora G. EGFR antagonists in cancer treatment. N Engl J Med 2008;358:1160–74.

17. Cappuzzo F, Ardizzoni A, Soto-Parra H, et al. Epidermal growth factor receptor targeted therapy by ZD 1839 (Iressa) in patients with brain metastases from non-small-cell lung cancer (NSCLC). Lung Cancer 2003;41:227–31.

18. Namba Y, Kijima T, Yokota S, et al. Gefitinib in patients with brain metastases from non-small-cell lung cancer: review of 15 clinical cases. Clin Lung Cancer 2004;6:123–8.

19. Mountain CF. Revisions in the International System for Staging Lung Cancer. Chest 1997;111:1710–7.

20. Tang X, Varella-Garcia M, Xavier AC, et al. EGFR abnormalities in the pathogenesis and progression of lung adenocarcinomas. Cancer Prev Res 2008;1: 404–8.

21. Hirsch FR, Varella-Garcia M, Bunn PA, Jr., et al. Epidermal growth factor receptor in non-small-cell lung carcinomas: correlation between gene copy number and protein expression and impact on prognosis. J Clin Oncol 2003;21:3798–807.

22. Merrick DT, Kittelson J, Winterhalder R, et al. Analysis of c-ErbB1/epidermal growth factor receptor and c-ErbB2/HER-2 expression in bronchial dysplasia: evaluation of potential targets for chemoprevention of lung cancer. Clin Cancer Res 2006;12:2281–8.

23. Tsao AS, Tang XM, Sabloff B, et al. Clinicopathologic characteristics of the EGFR gene mutation in non-small cell lung cancer. J Thorac Oncol 2006;1:231–9.

24. Varella-Garcia M. Stratification of non-small cell lung cancer patients for therapy with epidermal growth factor receptor inhibitors: the EGFR fluorescence *in situ* hybridization assay. Diagn Pathol 2006;1:19.

25. Massarelli E, Varella-Garcia M, Tang X, et al. KRAS mutation is an important predictor of resistance to therapy with epidermal growth factor receptor tyrosine kinase inhibitors in non-small-cell lung cancer. Clin Cancer Res 2007;13:2890–6.

26. Tang X, Shigematsu H, Bekele BN, et al. EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. Cancer Res 2005;65:7568–72.

27. Shigematsu H, Lin L, Takahashi T, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst 2005;97:339–46.

28. Rusch V, Klimstra D, Venkatraman E, Pisters PW, Langenfeld J, Dmitrovsky E. Overexpression of the epidermal growth factor receptor and its ligand transforming growth factor α is frequent in resectable non-small cell lung cancer but does not predict tumor progression. Clin Cancer Res 1997;3:515–22.

29. Volante M, Saviozzi S, Rapa I, et al. Epidermal growth factor ligand/receptor loop and downstream signaling activation pattern in completely resected nonsmall cell lung cancer. Cancer 2007;110:1321–8.

30. Zhang J, Iwanaga K, Choi KC, et al. Intratumoral epiregulin is a marker of advanced disease in non-small cell lung cancer patients and confers invasive properties on EGFR-mutant cells. Cancer Prev Res 2008;1:201–7.

31. Kakiuchi S, Daigo Y, Ishikawa N, et al. Prediction of sensitivity of advanced non-small cell lung cancers to gefitinib (Iressa, ZD1839). Hum Mol Genet 2004;13: 3029–43.

32. Wu W, O'Reilly MS, Langley RR, et al. Expression of epidermal growth factor (EGF)/transforming growth factor-α by human lung cancer cells determines their response to EGF receptor tyrosine kinase inhibition in the lungs of mice. Mol Cancer Ther 2007;6:2652–63.

33. Fernandes AM, Hamburger AW, Gerwin BI. Production of epidermal growth factor related ligands in tumorigenic and benign human lung epithelial cells. Cancer Lett 1999;142:55–63.

34. Gollamudi M, Nethery D, Liu J, Kern JA. Autocrine activation of ErbB2/ErbB3 receptor complex by NRG-1 in non-small cell lung cancer cell lines. Lung Cancer 2004;43:135–43.

35. Sasaki T, Nakamura T, Rebhun RB, et al. Modification of the primary tumor microenvironment by transforming growth factor α-epidermal growth factor receptor signaling promotes metastasis in an orthotopic colon cancer model. Am J Pathol 2008;173: 205–16.

36. Carpenter G. Nuclear localization and possible functions of receptor tyrosine kinases. Curr Opin Cell Biol 2003;15:143–8.

37. Johnson HM, Subramaniam PS, Olsnes S, Jans DA. Trafficking and signaling pathways of nuclear localizing protein ligands and their receptors. BioEssays 2004;26:993–1004.

38. Krolewski JJ. Cytokine and growth factor receptors in the nucleus: what's up with that? J Cell Biochem 2005;95:478–87.

39. Massie C, Mills IG. The developing role of receptors and adaptors. Nat Rev Cancer 2006;6:403 – 9.

40. Schlessinger J, Lemmon MA. Nuclear signaling by receptor tyrosine kinases: the first robin of spring. Cell 2006;127:45 – 8.

41. Fidler IJ. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. Nat Rev Cancer 2003;3:453 – 8.

42. Yonesaka K, Zejnullahu K, Lindeman N, et al. Autocrine production of amphiregulin predicts sensitivity to both gefitinib and cetuximab in EGFR wild-type cancers. Clin Cancer Res 2008;14:6963 – 73.

43. de Alava E, Ocana A, Abad M, et al. Neuregulin expression modulates clinical response to trastuzumab in patients with metastatic breast cancer. J Clin Oncol 2007;25:2656 – 63.

44. Cappuzzo F, Hirsch FR, Rossi E, et al. Epidermal growth factor receptor gene and protein and gefitinib sensitivity in non-small-cell lung cancer. J Natl Cancer Inst 2005;97:643 – 55.

45. Tsao MS, Sakurada A, Cutz JC, et al. Erlotinib in lung cancer—molecular and clinical predictors of outcome. N Engl J Med 2005;353:133 – 44.

46. Hirsch FR, Varella-Garcia M, McCoy J, et al. Increased epidermal growth factor receptor gene copy number detected by fluorescence *in situ* hybridization associates with increased sensitivity to gefitinib in patients with bronchioloalveolar carcinoma subtypes: a Southwest Oncology Group Study. J Clin Oncol 2005;28:6838 – 45.

47. Jackman DM, Holmes AJ, Lindeman N, et al. Response and resistance in a non-small-cell lung cancer patient with an epidermal growth factor receptor mutation and leptomeningeal metastases treated with high-dose gefitinib. J Clin Oncol 2006;24:4517 – 20.

48. Bunn PA, Jr., Dziadziuszko R, Varella-Garcia M, et al. Biological markers for non-small cell lung cancer patient selection for epidermal growth factor receptor tyrosine kinase inhibitor therapy. Clin Cancer Res 2006;12:3652 – 6.

49. Yatabe Y, Takahashi T, Mitsudomi T. Epidermal growth factor receptor gene amplification is acquired in association with tumor progression of EGFR-mutated lung cancer. Cancer Res 2008;68:2106 – 11.

50. Matsumoto S, Takahashi K, Iwakawa R, et al. Frequent EGFR mutations in brain metastases of lung adenocarcinoma. Int J Cancer 2006;119:1491 – 4.

Methylation of the candidate biomarker *TCF21* is very frequent across a spectrum of early stage non-small cell lung cancers

Kristy L. Richards[a,b,c], Baili Zhang[a], Menghong Sun[d], Wenli Dong[e], Jennifer Churchill[a], Linda L. Biachinski[a], Guosheng Yin[e], D. Neil Hayes[b], Ignacio I. Wistuba[d,f], Ralf Krahe[a,f,*]

[a]Department of Genetics, University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

Departments of [b]Medicine and [c]Genetics, The Lineberger Comprehensive Cancer Center, University of North Carolina School of Medicine, Chapel Hill, North Carolina, USA

Departments of [d]Pathology, [e]Biostatistics, and [f]Thoracic/Head and Neck Medical Oncology, University of Texas M. D. Anderson Cancer Center, Houston, TX, USA

*Corresponding author: Ralf Krahe, Department of Genetics, Unit 1010, University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030; USA; Phone: +1-713-834-6345; Fax: +1-713-834-6319; E-mail: rkrahe@mdanderson.org

Running Title: *TCF21* as a methylation biomarker in lung cancer

ABSTRACT

The transcription factor *TCF21* is involved in mesenchymal-to-epithelial differentiation and has been shown to be aberrantly hypermethylated in lung and head and neck cancers. Because of its reported high-frequency of hypermethylation in lung cancer, we sought to characterize the stages and types of non-small cell lung cancer (NSCLC) that are hypermethylated and to define the frequency of hypermethylation and associated "second hits". We determined *TCF21* promoter hypermethylation in 105 NSCLC including various stages and histologies in smokers and nonsmokers. Additionally, we examined *TCF21* loss-of-heterozygosity and mutational status. We also assayed 24 head and neck squamous cell carcinoma (HNSCC) samples and 22 cancer cell lines from varied tissue origins. We validated and expanded on our NSCLC results by examining TCF21 immunohistochemical expression on a tissue microarray containing 300 NSCLC cases. Overall, 81% of NSCLC samples showed *TCF21* promoter hypermethylation and 84% showed decreased TCF21 protein expression. Multivariate analysis showed that TCF21 expression, although below normal in both histologies, was lower in adenocarcinoma than squamous cell carcinoma, and was not independently correlated with gender, smoking and *EGFR* mutation status, or clinical outcome. HNSCC and cell lines from other cancer types also showed frequent *TCF21* promoter hypermethylation. Hypermethylation and decreased expression of *TCF21* were tumor-specific and very frequent in all NSCLC, even in early-stage, thus making *TCF21* a potential candidate methylation biomarker for early-stage NSCLC screening. TCF21 hypermethylation in HNSCC and a variety of tumor cell lines suggests it may also be a valuable methylation biomarker in other tumor types.

Keywords: TCF21; Methylation; Biomarker; Lung cancer; Screening

1.   Introduction

Lung cancer is the number one cause of cancer mortality worldwide, and kills more people than breast, colon, and prostate cancer combined.[1] Unlike these other common cancers, however, there is no effective screening strategy to detect early-stage lung cancer at a time when surgery may be curative. The need for such a strategy is obvious, and many attempts to detect lung cancer early have so far failed to show clinical benefit.[2-4] These include screening CT scans, sputum cytology, screening chest X-rays, serum markers.

Recently, promoter hypermethylation has been recognized as an important mechanism by which genes regulating cellular proliferation are silenced during cancer development[5,6] Promoter hypermethylation involves DNA methylation of CpG islands in or near the promoter region of certain genes, rendering them transcriptionally silent. This downregulation of gene expression of important cellular growth control genes has been shown to be important for cancer progression and outcome, with poorer outcomes associated with promoter hypermethylation of such important genes as *RASSF1A*, *RARB*, and *HIF1*.[7-9]

TCF21 is a recently recognized target of aberrant promoter hypermethylation in cancer, discovered in a genomic screen for regions of DNA that are hypermethylated in cancer.[10] It was reported to be frequently hypermethylated in head and neck and lung cancer, and restoration of *TCF21* expression inhibited tumor growth, both in a lung cancer cell line and in a mouse xenograft model. *TCF21* is widely expressed; its normal function is to promote mesenchymal transition into epithelial cells.[11] Reversal of this process, known as the epithelial-to-mesenchymal transition (EMT), has been implicated in tumor invasion and metastasis;[12,13] Therefore, silencing of *TCF21* may be a mechanism for tumor cells to gain these aggressive characteristics during the course of tumor progression.

Given that *TCF21* was reported to be frequently hypermethylated and silenced in NSCLC, as well as its plausible biologic role in tumor progression, we sought to more precisely define the frequency of *TCF21* promoter hypermethylation in NSCLC. We were especially

interested in defining its frequency among different cancer stages and histologic subtypes. Here, we show that *TCF21* is very frequently hypermethylated in a variety of NSCLC, and that protein expression of TCF21 is also very frequently reduced, either of which could be used for screening and/or diagnostic purposes as a biomarker of early disease.

2. Materials and Methods

2.1 Frozen tumor specimens, cell lines, and DNA extraction

Patient NSCLC specimens were obtained from surgical specimens at both the University of Texas M. D. Anderson Cancer Center (42 matched tumor/normal samples, 7 unpaired tumor samples) as well as from the University of North Carolina Lineberger Comprehensive Cancer Center tumor bank of surgical specimens (56 unpaired tumor samples). In both institutions, informed consent was obtained prior to surgery for the use of specimens as part of an IRB-approved protocol, in accord with the Helsinki Declaration. Tissue was snap-frozen and used for later DNA extraction. Genomic DNA was extracted from the DNA-protein phase of TriZol-extracted tissues according to the manufacturer's suggestions (Invitrogen). DNA was extracted using the PureGene kit (Gentra) on cell pellets from four HNSCC cell lines (SCC-4, SCC-9, SCC-15 and SCC-25), five lung cancer cell lines (H1395, H520, H2170, SK-MES-1 and SW-900), one breast cancer cell line (MCF7), one cervical cancer cell line (HeLa), two brain cancer cell lines (SK-N-AS and M059K), one uterine cancer cell line (AN3CA), one sarcoma cell line (HT1080), one kidney cancer cell line (HEK293), and six colon cancer cell lines (LoVo, SW48, HCT-15, DLD-1, COLO 320DM and RKO) according to the manufacturer's suggestions. All cell lines are available from ATCC (Manassas, VA). Four normal pools of DNA from peripheral blood mononuclear cells (PBMCs) were generated representing different genders and ages. Three of the pools (females ≤40 yrs of age, females age >40 yrs, and males ≤40 yrs) were comprised of five individuals per pool. The fourth pool (males ≥40 yrs) was comprised of six individuals.

2.2. *TCF21* promoter methylation

PCR and sequencing primers were designed using the PSQ Assay Design software (Qiagen). PCR was performed in a 25 µl reactions containing Qiagen HotStart Taq master mix (Qiagen) using 1 µl bisulfate-converted DNA (about 10 ng/µl). Bisulfite conversion of genomic DNA was

performed as previously reported.[14] Briefly, 0.5-1.0 µg of genomic DNA was treated using the

EZ-96 DNA Methylation Gold Kit (Zymo Research), including DNA sulfonation, deamination,

desalting, desulfonation and recovery. Bisulfite-treated DNA was stored at –80°C until use. To

reduce the cost per assay, an amplification protocol was developed using a biotinylated

universal primer approach.[14] Final primer concentrations were 10 nM of the reverse primer

tailed with the universal primer (5'-

GAC<u>GGGACACCGCTGATCGTTTA</u>CCAAAAAAAACCCCCTAA-3'), 100 nM of the untailed

forward primer (5'-GGTAGGGTGGTTTTGAGTT-3'), and 90 nM of the universal biotinylated

primer (5'-<u>GGGACACCGCTGATCGTTTA</u>-3') in each reaction. The universal primer sequence

is underlined. The predicted amplicon size was 153 bp. Amplification was carried out as

follows: denaturation at 95°C for 5 min, followed by 50 cycles at 95°C for 30 sec, 51°C for 1 min,

72°C for 45 sec, and a final extension at 72°C for 7 min.

Following PCR amplification, Pyrosequencing was performed on a PSQ96HS system

(Qiagen) according to the manufacturer's protocol including the use of single strand binding

protein (PyroGold reagents). The Pyrosequencing primer was (5'-TTGAGTTTGGAGAAGG-3').

The results were analyzed using Q-CpG software (Qiagen), which calculates the methylation

percentage ($^{m}C/(^{m}C+C)$) for each CpG site, allowing quantitative comparisons. The methylation

index (MI) was calculated as the average value of $^{m}C/(^{m}C+C)$ for all nine of the interrogated

CpG sites in the assay. Genomic DNA treated with M.*Sss*I (New England Biolabs) was used as

a universally methylated positive control; the same untreated genomic DNA amplified by whole

genome amplification (GenomiPhi, GE Healthcare) was used as a universally unmethylated

negative control.

2.3   LOH and mutation detection

Primers were designed for detection of four microsatellites within and flanking *TCF21*. Primer

sequences are shown in Supplemental Table S1. All forward primers were 5'-tailed with 5'-

GAC<u>GGGACACCGCTGATCGTTTA</u>-3' and all reverse primers were 5'-tailed with 5'-GTTTCTT-

3'. A universal primer with the sequence 5'-<u>GGGACACCGCTGATCGTTTA</u>-3' end-labeled with

either FAM, HEX, or NED was used in all microsatellite amplifications. PCR conditions for the

three primer reactions were as described above for amplification using the universal biotinylated

primer. Amplification products were pooled as appropriate and analyzed by capillary

electrophoresis on an ABI 3100 Genetic Analyzer (Applied Biosystems).

The coding region of *TCF21* (exons 1 and 2) was sequenced in both directions in four

fragments. In all, 45 lung cancer samples showing zero or one hit were sequenced. Samples

which had already been scored as having two hits were not sequenced. Primer sequences are

shown in Supplemental Table S1. All forward primers were 5'-tailed with M13 forward sequence

5-TGTAAAACGACGGCCAGT-3',and all reverse primers with M13 reverse 5'-

CAGGAAACAGCTATGACC-3'. After amplification, samples were treated with Exo-SAP

(Amersham), sequenced using Big Dye Terminator v3.1 (Applied Biosystems) under standard

conditions and products purified by ethanol precipitation, dehydrated in a vacuum centrifuge,

and resuspended in 20 µl formamide before capillary electrophoresis on an ABI 3100 Genetic

Analyzer. Sequences were aligned and visualized using Sequencer software (Gene Codes).

Fragment 1 contained a polymorphic $(CT)_n$ simple tandem repeat of 8 to 12 units, which, when

polymorphic, was used to confirm retention-of-heterozygosity identified by the microsatellites.

2.4   Archival NSCLC case selection and tissue microarray construction

We obtained archival, formalin-fixed and paraffin-embedded (FFPE) material from surgically

resected lung cancer specimens containing tumor and adjacent lung tissues from the Lung

Cancer Specialized Program of Research Excellence (SPORE) Tissue Bank at The University

of Texas M. D. Anderson Cancer Center, which was approved by the Institutional Review

Board. Tumor tissue specimens from 300 NSCLCs (191 adenocarcinomas, and 109 squamous

cell carcinomas) were histologically examined, classified using the 2004 World Health

Organization (WHO) classification system,[15] and selected for tissue microarray (TMA)

construction. After histologic examination, TMAs were constructed using triplicate 1-mm

diameter cores from each tumor. Detailed clinical and pathological information, including

demographic data, smoking history (never- and ever-smokers) and status (never, former, and

current smokers), pathologic TNM staging,[16] overall survival, and time of recurrence, was

available in most cases (Table 1). Patients who had smoked at least 100 cigarettes in their

lifetime were defined as smokers, and smokers who quit smoking at least 12 months before

lung cancer diagnosis were defined as former smokers.

2.5   Immunohistochemical staining and evaluation

An anti-human TCF21 antibody was used for immunostaining (ab32981, Abcam). FFPE tissue

histology sections (5-µm thick) were deparaffinized, hydrated, heated in a steamer for 10 min

with 10 mM sodium citrate (pH 6.0) for antigen retrieval. Peroxide blocking was performed with

3% H_2O_2 in methanol at room temperature for 15 min, followed by 10% bovine serum albumin in

TBS-t for 30 min. Slides were incubated with primary antibody at 1:200 dilution for 65 minutes

at room temperature. After washing with PBS, incubation with biotin-labelled secondary

antibody for 30 min followed. Finally, samples were incubated with a 1:40 solution of

streptavidin-peroxidase for 30 min. The staining was then developed with 0.05% 3',3-

diaminobenzidine tetrahydrochloride prepared in 0.05 mol/l Tris buffer at pH 7.6 containing

0.024% H_2O_2 and counterstained with hematoxylin. FFPE lung tissues having normal bronchial

epithelia were used as positive control. For a negative control, we used the same specimens

used for the positive controls, replacing the primary antibody with PBS.

TCF21 immunostaining was detected in the cytoplasm of epithelial and tumor cells.

Immunohistochemical expression was quantified by microscope observation by two pathologists

(M.S. and I.W.) using a four-value intensity score (0, 1+, 2+ and 3+) and the percentage of the

reactivity extent. A final score was obtained by multiplying both intensity and extension values

(range 0-300), and four levels of expression were arbitrarily calculated based on that score: (a)

negative (score 0-9); (b) low (score 10-100); (c) intermediate (score 100 to 199); and (d) and

high (score 200-300). Levels and scores were used for analysis.

2.6   *EGFR* mutation analysis

Exons 18 through 21 of *EGFR* were PCR amplified using intron-based primers as previously

described.[17,18] From microdissected FFPE cells, ~200 cells were used for each PCR

amplification. All PCR products were directly sequenced using the PRISM dye-terminator cycle

sequencing method (Applied Biosystems). All sequence variants were confirmed by

independent PCR amplifications from at least two independent microdissections and DNA

extraction, and sequenced in both directions, as previously reported.

2.7   Statistical analysis

Blip plots were generated to summarize the distribution of TCF21 protein expression levels from

the 300-sample TMA. Summary statistics of TCF21 protein expression levels for each

prognostic factor category were generated. Wilcoxon test, Kruskal-Wallis, and Fisher Exact test

were used to compare TCF21 expression among different levels of prognostic factors.

Spearman correlation coefficient was used to estimate the correlation between two continuous

variables (e.g., age and TCF21 expression). Kaplan-Meier survival curves for patient overall

survival and recurrence free survival were produced. The median survival times with the 95%

confidence intervals and the survival rates at 3 yrs and 5 yrs with the 95% confidence intervals

by prognostic factors were produced. To test the survival difference between the patient groups

for both overall and recurrence free survival, we used the log-rank test. Cox proportional hazard

models were utilized for multivariate analysis to examine the covariate effects. We constructed

the full model with all covariates included and the reduced model in which only significant

covariates and the TCF21 expression were included. The hazard ratios with the 95%

confidence intervals and the *p*-values were reported. All tests were two-sided.

3.   Results

3.1   *TCF21* is highly methylated in nearly all cancer cell lines

To characterize *TCF21* methylation levels in normal and malignant states, we examined various cancer cell lines from a spectrum of tissue types (brain, breast, cervix, colon, connective tissue, head and neck, kidney, lung, and uterus). We also assayed *TCF21* methylation in normal PBMCs from younger and older individuals of both sexes, since methylation levels can be influenced by age and/or sex. Universally methylated control DNA and genetically matched unmethylated control DNA defined the boundaries of detection of our assay (3-93% methylation). Using Pyrosequencing-based Methylation Analysis (PMA) we analyzed *TCF21* methylation by averaging methylation levels of nine promoter CpG sites. All but one cell line (SK-N-AS, a neuroblastoma cell line, 38%) was highly methylated, with levels at or approaching the upper limit of detection (Fig. 1). Normal PBMCs were essentially identical regardless of age or gender, and demonstrated moderate levels of baseline methylation at ~20%.

3.2   *TCF21* is hypermethylated in >80% of NSCLC

To define the threshold for hypermethylation positivity, we began our analysis using genetically matched NSCLC and adjacent normal tissue pairs from the same patient. To assess the baseline levels of *TCF21* methylation in lung tissue, we examined both normal adjacent tissue (NAT) from the tumor/normal (T/N) pairs (n=42) comparing them to PBMC. Average methylation levels in NAT were 21.5% (SD=4.6; n=42), and in the normal PBMC 20.1%. Average *TCF21* methylation levels in T samples were 41.3% (SD=11.6; n=42) (Fig. 2A). The difference between the average methylation levels in N and T tissues was highly significant (p-value<1×10^{-13}).

Using a threshold of 30% methylation, we found that 37 of 42 tumors (88%) were hypermethylated, while 41 of 42 matched normal samples (98%) were not. Using this cutoff to define hypermethylation, we then assayed second set of 63 unpaired NSCLC samples. This

second set of tumors contained a small number of large cell histologic subtypes, and some

mixed histologic types (mostly adeno-squamous). We found that 48 (76%) of them were

hypermethylated (Fig. 2B). Overall, the average methylation levels of all the tumor samples

combined was 39.2% (SD=11.7; n=105). Using the threshold of 30% methylation, the overall

frequency of hypermethylation in NSCLC was 81% (85/105).

3.3   *TCF21* is hypermethylated in >80% of HNSCC tumors

We also examined 27 HNSCC primary tumors. Similar to NSCLC, we found that average

methylation levels were 55.0% (SD=11.4; n=27) (Fig. 2B). Using the same threshold of 30%

methylation, 26 of 27 (96%) tumors showed *TCF21* hypermethylation. It should be noted,

however, that the appropriate cutoff for hypermethylation in HNSCC may be higher, if normal

adjacent tissue also has a higher baseline levels of methylation.

3.4   Reduced expression of TCF21 protein in NSCLC

To determine whether *TCF21* promoter hypermethylation also resulted in decreased TCF21

protein expression, we used a NSCLC TMA containing tumor samples from 300 patients. The

microarray was stained with a TCF21 antibody, and protein levels were scored as none, low,

intermediate, or high (Fig. 3A). While normal adjacent lung tissue stained strongly for TCF21,

253 of 300 (84%) NSCLC samples showed reduced (either low or none) staining (Fig. 3B).

Similar frequencies of *TCF21* hypermethylation and decreased protein expression

suggested that hypermethylation leads to reduced protein levels, which would be consistent with

previously reported decreased mRNA levels resulting from *TCF21* promoter hypermethylation.[10]

Because our TMA included only 9 overlapping samples between the TMA and *TCF21*

methylation sets, we assembled a smaller TMA with 31 samples overlapping (Supplemental

Table S2). Interestingly, *TCF21* hypermethylation and protein expression were independent

events (Supplemental Table S3B), suggesting that methods other than hypermethylation could

result in decreased protein expression.

3.5   *TCF21* **loss-of-heterozygosity and sequence analysis**

Because some NSCLC samples showed loss of TCF21 protein expression without

hypermethylation and the average levels of *TCF21* hypermethylation were ~40%, which might

not be expected to completely abolish protein expression, we examined potential "second hits"

at the *TCF21* locus (Supplemental Table S2). First, we examined loss-of-heterozygosity (LOH)

in 33 of the paired samples, using four microsatellite markers spanning the *TCF21* locus and

closely flanking region. LOH was seen in 14 (42%) of these samples, showing a trend toward

more frequent occurrence in samples that were not hypermethylated (*p*-value=0.09)

(Supplemental Table S3A). In addition to LOH, we sequenced the *TCF21* coding region in 45

lung cancer samples that showed either zero or one hit by methylation or LOH analysis.

Samples with both hypermethylation and LOH were not sequenced. No *TCF21* coding

mutations were found.

3.6   **Reduced TCF21 protein expression is widespread and independent of stage and other clinical features, but correlated with histology**

To determine whether TCF21 expression was correlated with clinical features such as gender,

race, stage, smoking status, histology, or prognosis, we performed univariate analysis (Table 1).

Histology and TCF21 expression showed the most significant correlation (*p*-value=0.0026);

however, smoking status (*p*-value=0.0209) and gender (*p*-value=0.0476) were also marginally

significant. Cox proportional hazards analysis was performed to assess association between

TCF21 and overall survival and recurrence-free survival, but neither association was significant,

in either a multivariate model (Table 1) or a univariate model (data not shown).

Given previously reported associations between smoking, gender, and histology with *EGFR* status,[19] we then analyzed the 202 patient subset for which *EGFR* status was known, for associations with TCF21 expression. In a multivariate analysis with histology, gender, and smoking status, only histology was statistically significantly (*p*-value = 0.007) associated with TCF21 levels, while smoking history and gender were not independently associated. When only adenocarcinomas were considered (n=172), *EGFR* status was not associated with TCF21 expression level, nor was *EGFR* status associated in a univariate analysis with all 202 patients. Therefore, the only significant correlation (*p*-value=0.007) is that adenocarcinomas have lower levels of TCF21 expression than SCCs, although all histologies have significantly lower TCF21 levels than those in normal tissue.

4.   Discussion

4.1   *TCF21* has the highest frequency of promoter hypermethylation in NSCLC of any gene known to date

Many genes have been reported to be hypermethylated in NSCLC.[20-24] However, the frequency of these events has not been high enough in all NSCLC subtypes for utilization as a screening tool, requiring combinations of genes to approach a sensitivity high enough for a screening test. Despite numerous reports of hypermethylated genes in NSCLC, identified by a variety of approaches, none has a reported frequency of hypermethylation as high as *TCF21*, except one that also examined *TCF21* itself, and a recent publication limited to only the SCC subtype of NSCLC.[20-23] This study was specifically focused on *TCF21* in NSCLC and the susceptibility locus at 6q23-q25. Among 43 genes selected in the region, *TCF21* had the highest rates of cancer-specific hypermethylation (81%),[23] exactly matching our rates of *TCF21* hypermethylation.

The high rates (80-85%) of *TCF21* promoter hypermethylation and decreased protein expression are high enough for *TCF21* to be used as a screening biomarker alone, either by increased methylation or decreased protein levels. The sensitivity of *TCF21* hypermethylation/decreased TCF21 protein expression compares favorably with that of prostate-specific antigen (PSA), the current screening biomarker for prostate cancer, which has been shown to be <4 (i.e., in the normal range) in 15% of men with prostate cancer, a sensitivity of 85%.[25] Of course, one of the main difficulties in lung cancer screening remains in the acquisition of relevant tissue (in this case early lung tumors), but detection of *TCF21* hypermethylation has been reported in biopsies and sputum samples, which is promising.[26]

4.2   Detection of *TCF21* hypermethylation by highly quantitative method

One significant advantage of methylation detection by Pyrosequencing-based Methylation Analysis (PMA) following bisulfite conversion is that quantitative levels can be measured across

multiple sites, rather than the more qualitative output obtained with methylation-specific PCR (MS-PCR) or other qualitative or semiquantitative methods (e.g., COBRA). PMA enabled us to reliably detect a difference between the 20% average methylation in N tissue, and 40% average methylation in T tissue. This difference would likely not have been detected with less quantitative methylation detection strategies. It is possible that other genes known to be hypermethylated in NSCLC may prove to be more sensitive and/or specific, if more quantitative methods similar to Pyrosequencing were to be routinely applied. The 40% methylation levels in NSCLC tissue raises the question of whether only one of the two *TCF21* alleles is silenced by hypermethylation, or whether 40% of cells have both alleles silenced, either of which could produce the observed result. It is interesting that hypermethylation of 40% of alleles is frequently associated with completely absent TCF21 protein expression, suggesting either that the second allele is silenced by a different mechanism than hypermethylation, or that there is a threshold level of gene expression necessary to produce detectable TCF21 protein levels.

4.3   Reduction of TCF21 protein levels similar to *TCF21* hypermethylation rates

In addition to *TCF21* hypermethylation, we also examined the downstream effect of this hypermethylation by examining protein expression directly. In both cases we found *TCF21* hypermethylation/decreased TCF21 protein levels at similar rates--81% and 84%, respectively. Given that decreased mRNA expression of *TCF21* has been shown to result from promoter hypermethylation,[10] the similar rates of hypermethylation and decreased protein expression are consistent with the notion that decreased mRNA expression results in decreased protein expression. However, since there were cases with low/absent protein expression despite normal *TCF21* methylation levels, other regulatory mechanisms likely are in effect. LOH occurs at a rate of 42%. Since LOH occurs in at least a few cases without *TCF21* hypermethylation, this implies inactivation of *TCF21* in other ways. Since we did not detect any coding mutations, these could be promoter or other regulatory region DNA mutations. Alternatively, dysregulation

by micro-RNAs could be a factor. Interestingly, the sole predicted regulator of *TCF21* is miR-92a,[27] which is overexpressed in a variety of cancers.[28,29]

4.4 TCF21 is an excellent candidate biomarker for early lung cancer detection

Several characteristics of *TCF21* make it an attractive target for screening efforts in NSCLC.

First, it is hypermethylated at similar frequencies in all histologic subtypes of NSCLC examined, including early- and late-stage cancers. Second, it has a higher frequency of hypermethylation than any gene published to date in NSCLC, without subdivision by histologic subtype.[10,20-23] This high sensitivity is combined with a high specificity as well. We detected a false-positive rate of only 1 in 42 samples with NAT, for a specificity of 98%. In other reported control tissues, such as PBMCs and human bronchial epithelial cells (HBECs) from smokers, there were no false-positives (n=20 in each case).[23] The high specificity in normal adjacent tissue is especially noteworthy in that there appears to be no evidence for a "field-effect", which can complicate screening in smokers who often have cancers arising in a field of premalignant lesions, leading to false-positive screening results. Instead, the very low prevalence of *TCF21* hypermethylation in NAT that we report suggests that *TCF21* hyermethylation is restricted to cancerous tissue only.

In summary, we have established that *TCF21* hypermethylation and reduced TCF21 protein are ubiquitous in NSCLC, occurring in 80-85% of tumors across a wide variety of stages, histologies, and other clinical characteristics. Given the high rate of increased methylation and decreased protein expression, combined with their lack in normal adjacent tissue, we propose that *TCF21* is an excellent biomarker for further development as a lung cancer screening tool.

Acknowledgements

Conflict of Interest: None declared.

REFERENCES

1.   Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005;**55**:74-108.

2.   Humphrey LL, Teutsch S, Johnson M. Lung cancer screening with sputum cytologic examination, chest radiography, and computed tomography: an update for the U.S. Preventive Services Task Force. *Ann Intern Med* 2004;**140**:740-53.

3.   Rossi A, Maione P, Colantuoni G, Gaizo FD, Guerriero C, Nicolella D, et al. Screening for lung cancer: New horizons? *Crit Rev Oncol Hematol* 2005;**56**:311-20.

4.   Chorostowska-Wynimko J, Szpechcinski A. The impact of genetic markers on the diagnosis of lung cancer: a current perspective. *J Thorac Oncol* 2007;**2**:1044-51.

5.   Vucic EA, Brown CJ, Lam WL. Epigenetics of cancer progression. *Pharmacogenomics* 2008;**9**:215-34.

6.   Duffy MJ, Napieralski R, Martens JW, Span PN, Spyratos F, Sweep FC, et al. Methylated genes as new cancer biomarkers. *Eur J Cancer* 2009;**45**:335-46.

7.   Kim YT, Park SJ, Lee SH, Kang HJ, Hahn S, Kang CH, et al. Prognostic implication of aberrant promoter hypermethylation of CpG islands in adenocarcinoma of the lung. *J Thorac Cardiovasc Surg* 2005;**130**:1378.

8.   Maruyama R, Sugio K, Yoshino I, Maehara Y, Gazdar AF. Hypermethylation of FHIT as a prognostic marker in nonsmall cell lung carcinoma. *Cancer* 2004;**100**:1472-7.

9.   Safar AM, Spencer H, 3rd, Su X, Coffey M, Cooney CA, Ratnasinghe LD, et al. Methylation profiling of archived non-small cell lung cancer: a promising prognostic system. *Clin Cancer Res* 2005;**11**:4400-5.

10.   Smith LT, Lin M, Brena RM, Lang JC, Schuller DE, Otterson GA, et al. Epigenetic regulation of the tumor suppressor gene TCF21 on 6q23-q24 in lung and head and neck cancer. *Proc Natl Acad Sci U S A* 2006;**103**:982-7.

11.   Quaggin SE, Schwartz L, Cui S, Igarashi P, Deimling J, Post M, et al. The basic-helix-loop-helix protein pod1 is critically important for kidney and lung organogenesis. *Development* 1999;**126**:5771-83.

12.   Baum B, Settleman J, Quinlan MP. Transitions between epithelial and mesenchymal states in development and disease. *Semin Cell Dev Biol* 2008;**19**:294-308.

13.   Lee JM, Dedhar S, Kalluri R, Thompson EW. The epithelial-mesenchymal transition: new insights in signaling, development, and disease. *J Cell Biol* 2006;**172**:973-81.

14.   Colella S, Shen L, Baggerly KA, Issa JP, Krahe R. Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. *Biotechniques* 2003;**35**:146-50.

15.   Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC. Tumours of the lung. In: Travis WD, Brambilla E, Muller-Hermelink HK, Harris CC, editors. Pathology and Genetics: Tumours of the Lung, Pleura, Thymus and Heart, World Health Orgnaziation Classification of Tumours. . Lyon: International Agency for Research on Cancer (IARC); 2004. p. 9-124.

16.   Mountain CF. Revisions in the International System for Staging Lung Cancer. *Chest* 1997;**111**:1710-7.

17.   Shigematsu H, Lin L, Takahashi T, Nomura M, Suzuki M, Wistuba, II, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. *J Natl Cancer Inst* 2005;**97**:339-46.

18.   Tang X, Shigematsu H, Bekele BN, Roth JA, Minna JD, Hong WK, et al. EGFR tyrosine kinase domain mutations are detected in histologically normal respiratory epithelium in lung cancer patients. *Cancer Res* 2005;**65**:7568-72.

19.   Kris MG, Natale RB, Herbst RS, Lynch TJ, Jr., Prager D, Belani CP, et al. Efficacy of gefitinib, an inhibitor of the epidermal growth factor receptor tyrosine kinase, in symptomatic patients with non-small cell lung cancer: a randomized trial. *JAMA* 2003;**290**:2149-58.

20. Esteller M, Corn PG, Baylin SB, Herman JG. A gene hypermethylation profile of human cancer. *Cancer Res* 2001;**61**:3225-9.

21. Belinsky SA. Gene-promoter hypermethylation as a biomarker in lung cancer. *Nat Rev Cancer* 2004;**4**:707-17.

22. Hoque MO, Kim MS, Ostrow KL, Liu J, Wisman GB, Park HL, et al. Genome-wide promoter analysis uncovers portions of the cancer methylome. *Cancer Res* 2008;**68**:2661-70.

23. Tessema M, Willink R, Do K, Yu YY, Yu W, Machida EO, et al. Promoter methylation of genes in and around the candidate lung cancer susceptibility locus 6q23-25. *Cancer Res* 2008;**68**:1707-14.

24. Anglim PP, Galler JS, Koss MN, Hagen JA, Turla S, Campan M, et al. Identification of a panel of sensitive and specific DNA methylation markers for squamous cell lung cancer. *Mol Cancer* 2008;**7**:62.

25. Thompson IM, Pauler DK, Goodman PJ, Tangen CM, Lucia MS, Parnes HL, et al. Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. *N Engl J Med* 2004;**350**:2239-46.

26. Shivapurkar N, Stastny V, Xie Y, Prinsen C, Frenkel E, Czerniak B, et al. Differential methylation of a short CpG-rich sequence within exon 1 of TCF21 gene: a promising cancer biomarker assay. *Cancer Epidemiol Biomarkers Prev* 2008;**17**:995-1000.

27. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res* 2006;**34**:D140-4.

28. Matsubara H, Takeuchi T, Nishikawa E, Yanagisawa K, Hayashita Y, Ebi H, et al. Apoptosis induction by antisense oligonucleotides against miR-17-5p and miR-20a in lung cancers overexpressing miR-17-92. *Oncogene* 2007;**26**:6099-105.

29.   Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, et al. A microRNA

        expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad*

   *Sci U S A* 2006;**103**:2257-61.

Figure Legends

Fig. 1 -- Methylation levels of individual cancer cell lines, normal PBMCs, and positive and negative methylation controls. *TCF21* promoter methylation levels are shown for 22 cancer cell lines, four pools of PBMCs of different sex (male, female) and age (≤40, >40 yrs). Control samples, fully methylated by treatment with *Sss*I methylase or fully unmethylated by whole genome amplification, are also shown.

Fig. 2 -- Methylation levels and percent of tumors with methylation levels >30% threshold. **(A)** *TCF21* promoter methylation levels are shown in a box and whisker plot for 42 normal adjacent lung tissues, 42 NSCLC tumors, 63 additional NSCLC tumors, all 105 NSCLC tumors combined, and 24 HNSCC tumors. **(B)** Bar graph representing the number of NSCLC and HNSCC tumors exceeding the 30% threshold for hypermethylation.

Fig. 3 -- TCF21 protein expression. **(A)** TCF21 immunohistochemical expression in lung cancer. NSCLC samples were stained with an anti-TCF21 antibody and scored as none, low, medium, or high. Representative examples of SCC and adenocarcinoma samples with high TCF21 expression (left) and no TCF21 staining (right) are shown. **(B)** Frequency of TCF21 expression in NSCLC on the TMA (300 patients). The percentage of samples in each expression category is shown. Reduced expression was defined as "none" or "low".

The y-axis is labeled "*TCF21* methylation level (%)" with values from 0.00 to 100.00 in increments of 10.00.

The x-axis is labeled "Cell Line" with categories: HEK-293, H520, H1395, H2170, AN3 CA, COLO 320DM, DLD-1, HeLa, HT-1080, HCT-15, LoVo, MCF7, M059K, RKO, SCC-15, SCC-25, SCC-4, SCC-9, SK-N-AS, SK-MES-1, SW48, SW900, NP F<40, NP F>40, NP M<40, NP M>40, Unmeth., Meth.

A

B

A

High   None

SCC

AC
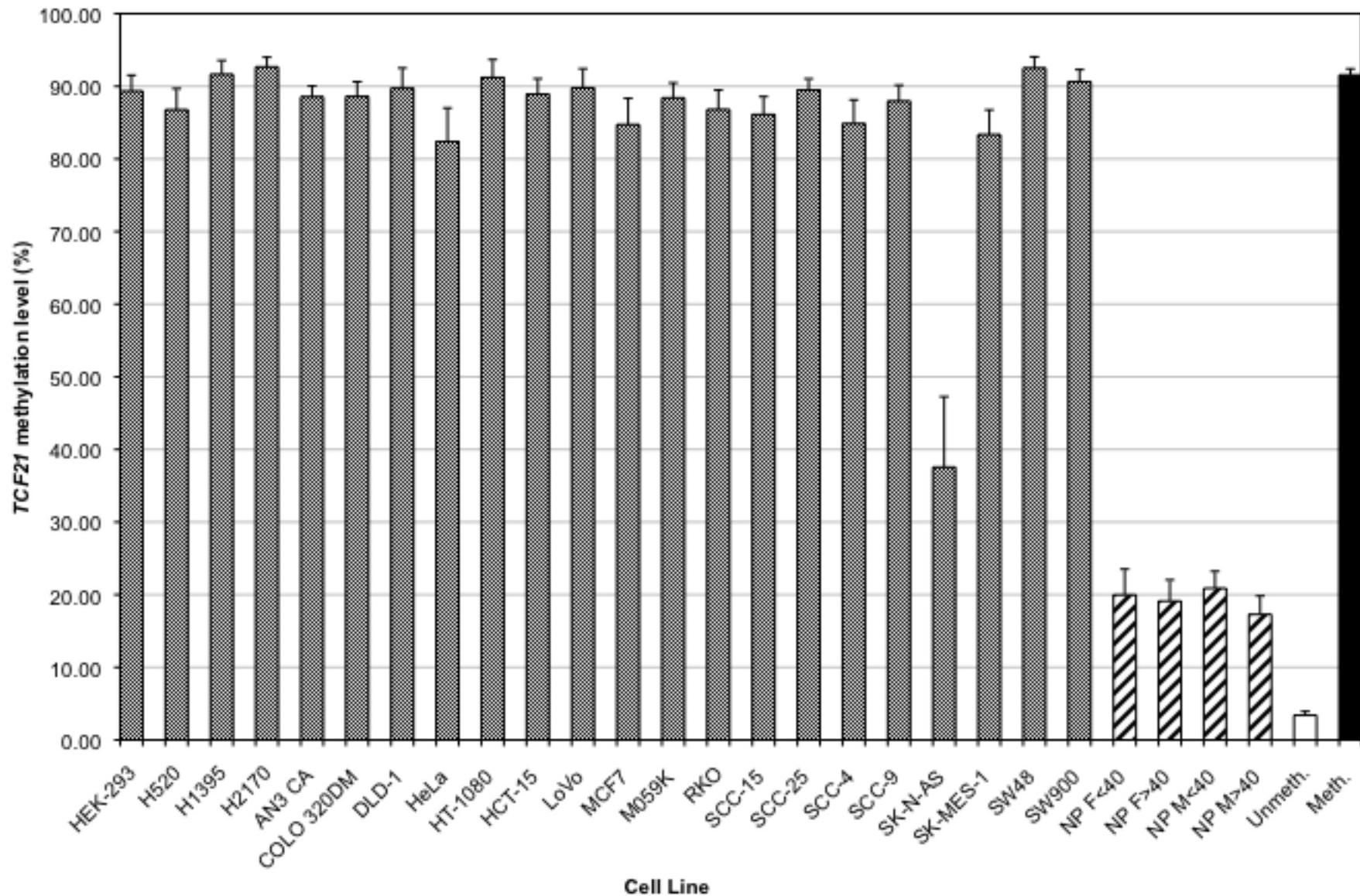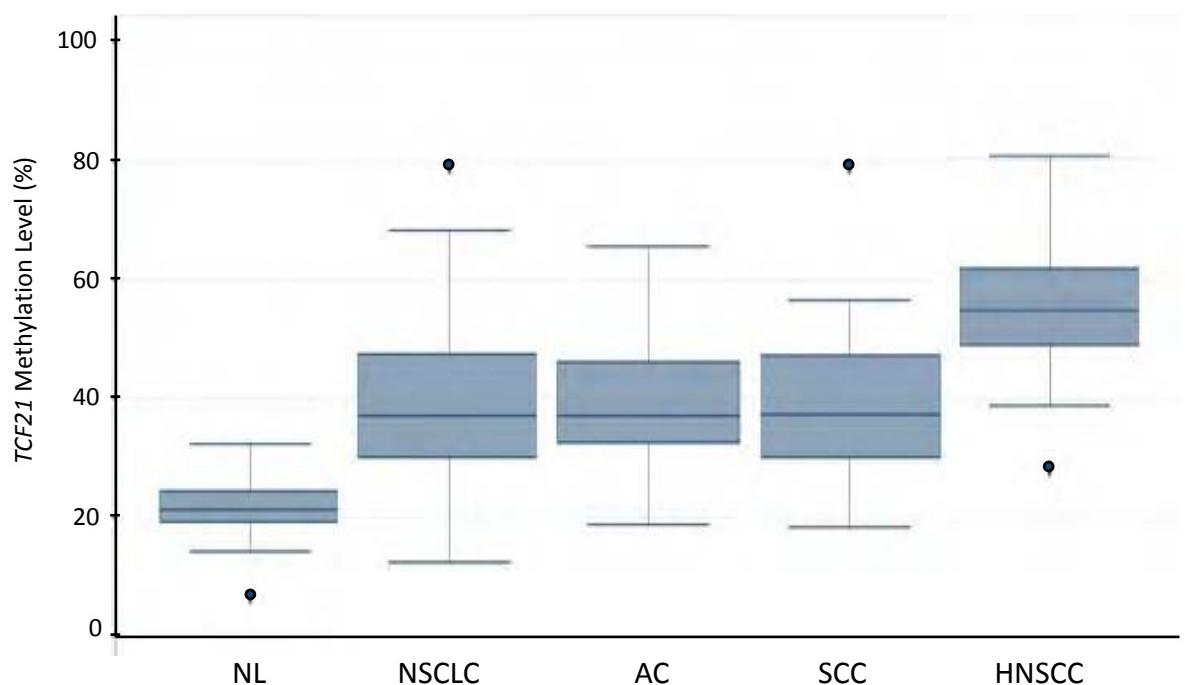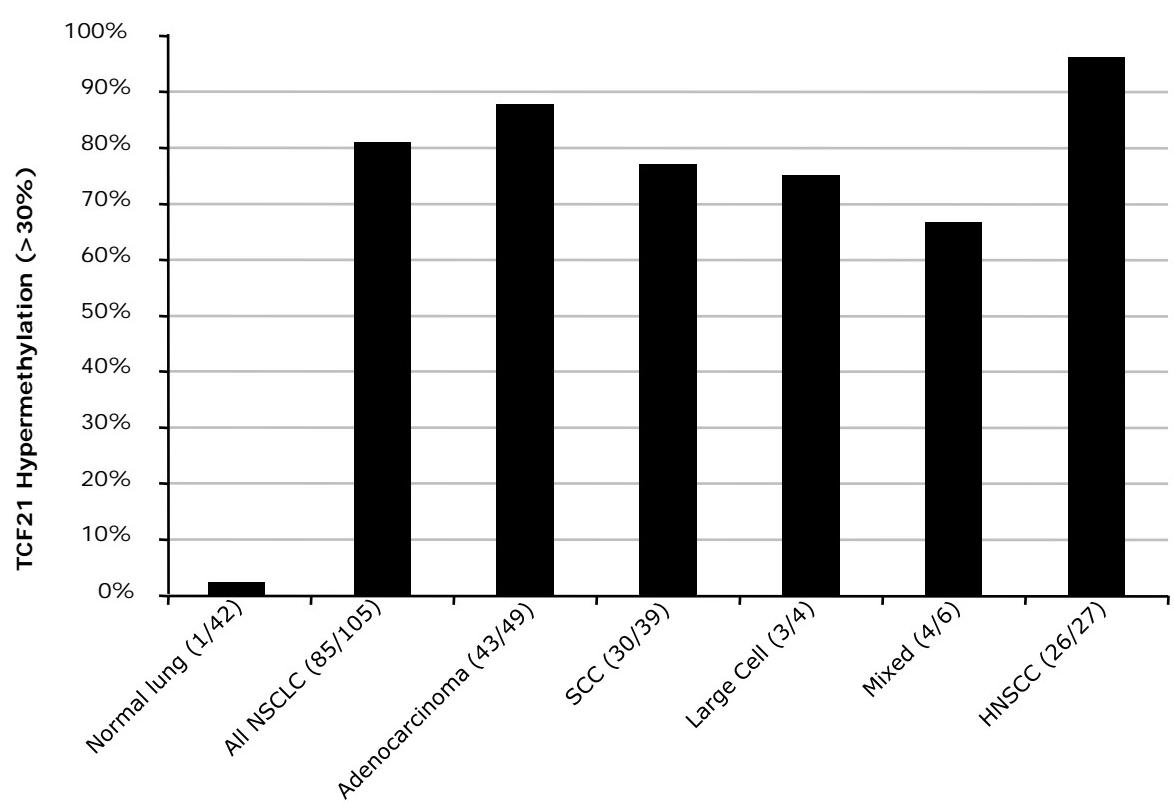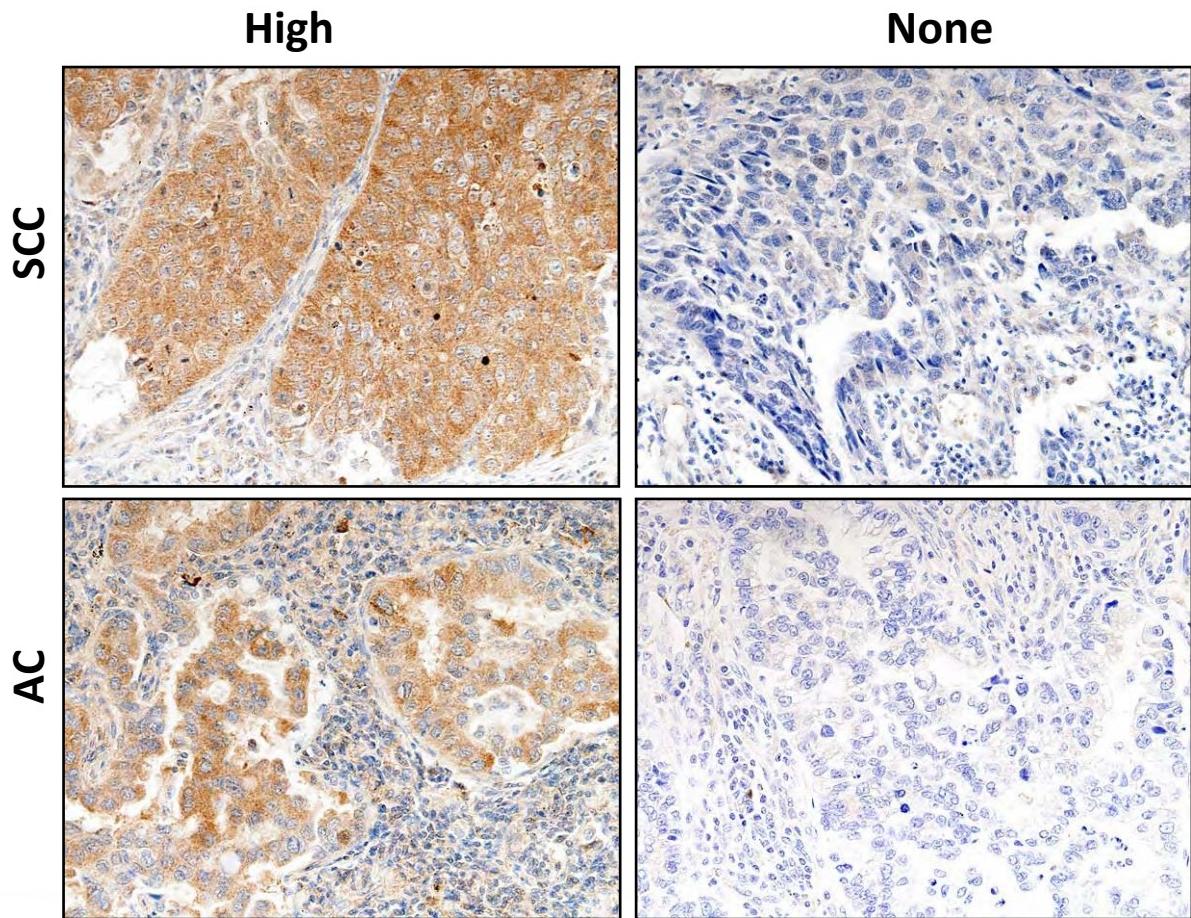
B
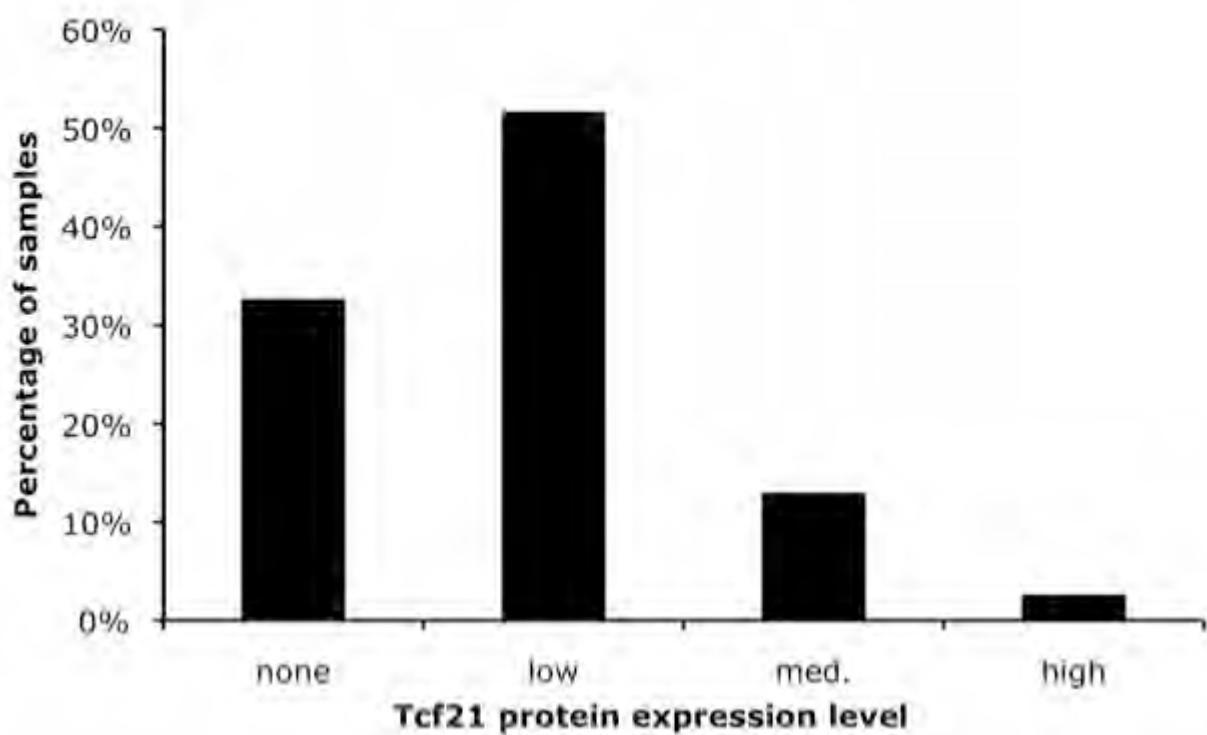
Supplemental Table S1 -- Primers used for LOH and mutation detection in TCF21.

| Analysis | Forward Primer | Reverse Primer | Amplicon |
LOH	(5'—3')	(5'—3')	Size (bp)
TCF21_5'_GT19	CATGCCTGGGCGACAG	GGCCCTTAAGCTGACAACTT	131
TCF21_CA14	ATGCAGTAAGGCCATAGTTTG	CAGCAGTGAGAGAACACCC	294
TCF21_E3_CT13	GTGTTTCAAGTAAGCGAGTCT	GGGAAAGCCTGCTAGAAT	163
TCF21_3'_GT12-AT8	AAGCTTAAGACTTTGGCTAA	CTAGTCTGGTTCCCTAGAGTT	208
Sequence			
TCF21_E1-1	ATTGAGTTTCCCTCCGGTTGTGAA	CTCGCAGTTGGAGCTCTCCTCGGTG	547
TCF21_E1-2	TCAGCGATGTGGAGGACCTTCAAGA	GCGGTGGTCGAGATGTGTAAGTCA	563
TCF21_E2-1	CCCCCTTCCTTTCATCTCAG	AATTACATATTGCACTTGGACCAGC	449
TCF21_E2-2	AAGATTCCCATCTATTTAACTTTA	ATCTGCATCTTCATTATGAAACTCA	419

Supplemental Table S2 -- Methylation, LOH, and protein expression data for TCF21 in lung cancer samples.

Sample ID	Histology	Methylation (%)	Hypermethylated (yes/no)	LOH	No. of Hits	TCF21 protein expression on TMA
462	Adeno	37.21%	yes	nd		negative
612	Adeno	32.74%	yes	nd		high
645	Adeno	31.39%	yes	nd		negative
649	Adeno	39.24%	yes	ROH	1	low
756	SCC	39.34%	yes	ROH	1	negative
758	SCC	43.11%	yes	nd		intermediate
759	Adeno	35.89%	yes	nd		low
782	Adeno	30.23%	yes	ROH	1	intermediate
794	SCC	51.55%	yes	LOH	2	negative
798	Adeno	33.43%	yes	LOH	2	low
799	SCC	54.54%	yes	LOH	2	intermediate
801	Adeno	51.64%	yes	nd		low
840	Adeno	35.06%	yes	LOH	2	intermediate
842	Adeno	28.46%	no	nd		low
846	Adeno	22.59%	no	LOH	1	negative
848	Adeno	44.62%	yes	LOH	2	intermediate
870	SCC	20.62%	no	nd		high
879	SCC	79.12%	yes	ROH	1	intermediate
913	SCC	59.04%	yes	ROH	1	intermediate
931	SCC	48.34%	yes	LOH	2	low
1114	Adeno	62.16%	yes	ROH	1	nd
1140	SCC	40.10%	yes	ROH	1	intermediate
1147	Adeno	56.92%	yes	LOH	2	low
1153	Adeno	38.89%	yes	ROH	1	low
1200	SCC	47.69%	yes	ROH	1	negative
1234	SCC	40.23%	yes	ROH	1	low
1237	SCC	39.78%	yes	LOH	2	intermediate
1244	SCC	42.65%	yes	LOH	2	low
1274	Adeno	44.23%	yes	ROH	1	nd
1289	Adeno	39.44%	yes	ROH	1	low
1323	Adeno	33.34%	yes	ROH	1	nd
1352	Adeno	65.74%	yes	LOH	2	nd
1355	Adeno	43.28%	yes	ROH	1	nd
1406	Adeno	50.94%	yes	ROH	1	nd
1416	Adeno	46.27%	yes	LOH	2	nd
1436	Adeno	30.04%	yes	ROH	1	nd
1453	SCC	45.70%	yes	ROH	1	nd
771_3	SCC	18.20%	no	LOH	1	intermediate
786_2	SCC	29.73%	yes	LOH	2	low
793_2	SCC	22.67%	no	nd		intermediate
878_2	SCC	33.74%	yes	ROH	1	nd
905_1	SCC	31.90%	yes	ROH	1	nd

Adeno, adenocarcinoma; SS, squamous cell carcinoma; nd, not done; ROH, retention-of-heterozygosity; LOH, loss-of-heterozygosity; No. of Hits, total of hypermethylation and LOH events (if both were assayed): 0 is neither, 1 is either hypermethylation or LOH, and 2 is both; TMA, tissue microarray.

Supplemental Table S3 -- Chi-squared test for TCF21hypermethylation coincidence with either TCF21 LOH (A) or TCF21 protein expression (B).

A. TCF21 Hypermethylation and LOH

observed

	hypermethylation	no hypermethylation
LOH	12	2
ROH	19	0

expected

	hypermethylation	no hypermethylation
LOH	13.15	0.85
ROH	17.85	1.15

p-value=0.1723

B. TCF21 Hypermethylation and protein expression

observed

	hypermethylation	no hypermethylation
negative	5	1
low	11	1
intermediate	9	2
high	1	1

expected

	hypermethylation	no hypermethylation
negative	5.03	0.97
low	10.06	1.94
intermediate	9.23	1.77
high	1.68	0.32

p-value=0.4075

TITF-1 and EGFR gene copy variations are associated with prognosis for the patients with non-small cell lung cancer

Ximing Tan, Diane Liu, Carmen Behrens, Dandan He, Menghong Sun, David Rice, J. Jack Lee, Waun K. Hong, and Ignacio I. Wistuba

UT MD Anderson Cancer Center, Houston, TX 77030

Thyroid transcription factor -1 (TITF-1, a lineage-specific transcription factor), and the epidermal growth factor receptor (EGFR, a tyrosine kinase membrane receptor) have shown frequent gene amplification in non-small cell lung carcinomas (NSCLC). We investigated the clinico-pathologic characteristics of NSCLCs having *TITF-1* andor *EGFR* gene copy number abnormalities by examining gene copy number status using quantitative polymerase chain reaction (qPCR) and DNA extracted from microdissected formalin-fixed and paraffin-embedded tissue from 53 adenocarcinomas and 29 squamous cell carcinomas. ß-actin gene was used as reference. In tumors, gene copy ratio referred to *ß-actin* ranged from 0.22 to 74.93 (median=1.52) for *TITF-1*, and 0.05 to 6.28 (median=1.51) for *EGFR*. Ratios 1 to 2 were defined as normal gene copy number (NGC). Ratios <1 and >2 were defined as low gene copy (LGC) and high gene copy (HGC) number, respectively. Both, LGC and HGC categories were defined as abnormal gene copy. Similar frequencies of *TITF-1* and *EGFR* copy number categories were detected comparing adenocarcinoma (*TITF-1*: LGC 15, 28%; NGC 20, 38%; HGC 18, 34%; *EGFR*: LGC 9, 17%; NGC 27, 51%; HGC, 17, 32%) with squamous cell carcinoma (*TITF-1*: LGC 8, 28%; NGC 14, 48%; HGC 7, 24%; *EGFR* : LGC 5, 17%; NGC 18, 62%; HGC 6, 21%). We found a statistically significant correlation between *TITF-1* and *EGFR* copy numbers (Spearman correlation coefficient=0.36, *P*=0.0008). In both tumor histologies, neither *TITF-1* nor *EGFR* gene copy increase (ratio >2) correlated with disease prognosis**.** However, in adenocarcinomas, Kaplan-Meier and log rank tests revealed that the median time to death was longer in patients with normal copy number compared with those with abnormal copies for *TITF-1* (median 4.76 years, 95% CI 2.95~NA, *P* = 0.04) and *EGFR* (4.76 years, 95% CI 3.13~ NA, *P*=0.04). Moreover, adenocarcinoma patients with combined *TITF-1* and *EGFR* abnormal copy showed worse overall survival (3.56 years, 95% CI 3.13~ NA) compared with patients with normal copy

status (median not reached, *P*=0.003). In these patients, multicovariate Cox modeling indicated that combined copy abnormality of both genes is an independent factor for worse overall survival (HR 4.566, *P*=0.0057). Our findings suggest that loss and gain of *TITF1* and *EGFR* are frequent abnormalities in both adenocarcinomas and squamous cell carcinomas of the lung, and in adenocarcinoma patients correlate with disease outcome. (Supported by Grant DoD-W81XWH-04-1-0142 and W81XWH-05-2-0027).

(*TITF-1* and *EGFR* gene copies in lung cancer)

TITF-1 Protein Expression Associates with Gene Methylation and Gene Copy Gain in Non-Small Cell Lung Carcinoma.

Ximing Tang, Fei Yang, Jianan Huang, Denise Woods, Alejandro Corvalan, Ignacio I. Wistuba

UT MD Anderson Cancer Center, Houston, TX 77030, USA

Background: The thyroid transcription factor-1 (TITF-1, NKX2-1) is a lineage-specific transcription factor normally expressed in peripheral pulmonary epithelial cells. TITF-1 protein is frequently expressed in adenocarcinoma, and absent in squamous cell carcinoma of the lung. *TITF-1* gene copy number gain (CNG) has been reported in a relatively small subset of adenocarcinoma of the lung, and we have also recently detected this abnormality in lung squamous cell carcinomas (Tang et al, unpublished data). To better characterize the mechanisms responsible to the different pattern of expression of TITF-1 protein in these two types of non-small cell lung carcinomas (NSCLC), we investigated the correlation between protein expression with CNG and methylation abnormalities in a large series of NSCLC specimens.

Methods: DNA extracted from NSCLC cell lines (n=21) and tumor tissue (n=173) samples was analyzed for *TITF*-1 gene methylation by bisulfite DNA sequencing, CNG by quantitative-PCR (qPCR), and protein expression by immunohistochemistry (IHC).

Results: 14 out of 21 NSCLC cell lines demonstrated *TITF-1* gene methylation in exons 1 and 2, but not in the promoter region. In the 14 methylated cell lines, the percentage of average methylation of CpG sites in exon 2 (43.9%) was much higher than in exon 1 (6.7%). Methylation was more frequently detected in squamous cell carcinoma (48/72, 72%) compared with adenocarcinoma (46/106, 43%, P=0.002). *TITF-1* CNG was detected in 35% of adenocarcinoma and 25% of squamous cell carcinoma, respectively. In adenocarcinomas, TITF-1 protein overexpression was associated with CNG; while in both tumor histologies reduced expression of the protein was correlated with gene methylation (R=-0.4, P=0.002). In addition, nearly 50% of CNG NSCLCs were detected with the gene methylation simultaneously, which indicated two opposite modulation mechanisms can exist in some NSCLC. No gene methylation and CNG were detected in non-malignant lung tumor tissues adjacent to lung tumors having those abnormalities. Conclusion: Our findings

indicate that, in NSCLC, TITF-1 increase and lack/reduction of protein expression correlates with gene copy number gain and methylation of the gene, respectively, suggesting that these two mechanisms are important for TITF-1 protein expression modulation in lung tumors.

Analysis of Copy Number Gain of *VEGF and VEGFR2/KDR* in Non-Small Cell Carcinoma (NSCLC) Cancer Cell Lines and Primary Tumors

Fei Yang[1], Ximing Tang[1], Alejandro Corvalan[1], Carmen Behrens[1], Heather Lin[1], Maria Gabriela Raso[1], J. Jack Lee[1], John D. Minna[2], and Ignacio Ivan Wistuba[1]

University of Texas MD Anderson Cancer Center, Houston, TX[1], and University of Texas Southwestern Medical Center, Dallas, TX[2].

Background. The VEGF family plays an important role in lung cancer angiogenesis. The VEGF receptor 2 (VEGFR2) is the predominant mediator of VEGF-stimulated endothelial cell migration, proliferation, survival and enhanced vascular permeability. Recently, *VEGF* copy number gain, and *VEGFR2/KDFR* copy number gain and mutation, have been described in lung adenocarcinoma tissue specimens. To better characterize the molecular changes of these two genes in NSCLC, we investigated their frequency of abnormalities in NSCLC cell lines and tumor tissue specimens.

Methods. We studied DNA extracted from 37 NSCLC cell lines and 93 tumor tissues obtained from surgically resected tumors, including 56 adenocarcinomas and 37 squamous cell carcinomas. *VEGF* and *VEGFR2/KDR* copy number were examined by quantitative (q)-PCR, and *VEGFR2/KDR* mutations of exons 7, 11, 21, 26, 27 and 30, by PCR-based sequencing. Protein expression of VEGF, VEGFR2 and CD34 was detected by immunohistochemistry in 45 lung cancer patients who had not received neoadjuvant chemotherapy.

Results. *VEGF* and *VEGFR2/KDR* copy number gains were detected at low frequencies in NSCLC cell lines and tumor tissue specimens. In cell lines, the frequency of copy number gains was 1/37 (copy number=4.6) for *VEGF*, and 1/37 (copy number=5.0) for *VEGFR2/KDR*. In primary tumors, copy number gains were detected in 3/93 (3.2%; 1 adenocarcinoma, and 2 squamous cell carcinomas) for *VEGF*, and 11/93 (11.8%; 4 adenocarcinomas, and 7 squamous cell carcinomas) for *VEGFR2/KDR*. These copy number changes were not detected in non-malignant lung tissues adjacent to the tumors in the 15 lung cancers with or without copy number increase of *VEGFR2/KDR*. Significant increase of cytoplasmic VEGFR2 expression was detected in lung cancer cases with *VEGFR2/KDR* copy number gain (P=0.004). There is no correlation between *VEGFR2/KDR* copy number gain with MVD-CD34 or VEGF expression in lung cancer. Mutation of *VEGFR2/KDR* was detected only in lung cancer cell line HCC2279 (CGT946CAT). No mutation of *VEGFR2/KDR* was found in primary lung tumors. 3 Single nucleotide polymorphisms were detected in *VEGFR2/KDR* exon 7, 11 and 21 in different frequency in primary lung cancer.

Conclusion. Our findings indicate that copy number gain of *VEGF* and *VEGFR2/KDR* are identified at low frequencies in NSCLC, affecting to both major histology types, adenocarcinoma and squamous cell carcinoma. A better characterization of the clinicopathologic features of the tumors having such genetic abnormalities will help to identify their role in lung cancer pathogenesis and progression, and eventually in the response to VEGFR targeted therapy. (Supported in part by grants by Grants DoD-W81XWH-04-1-0142 and US DoD W81XWH-07-1-0306).)